DOCUMENT RESUME

ED 028 999 24 SP 002 633

By-Sharpe, Donald M.

Isolating Relevant Variables in Student Teaching Assessment. Final Report.

Indiana State Univ., Terre Haute. School of Education.
Spons Agency-Office of Education (DHEW), Washington, D.C. Bureau of Research.
Bureau No-BR-6-1321
Pub Date Jan 69
Contract-OEC-3-7-061321-0342
Note-171p.
EDRS Price MF-$0.75 HC-$8.65
Descriptors-*Lesson Observation Criteria, *Measurement Instruments, *Predictive Validity, *Reliability, Secondary School Teachers, Student Teachers, Teacher Behavior
Identifiers-Secondary Student Teacher Performance Profile, SSTPP, TCAP, Teacher Classroom Activity Profile

Two instruments for evaluating the classroom behavior of secondary student teachers, the Secondary Student Teacher Performance Profile (SSTPP) and the Teacher Classroom Activity Profile (TCAP), were developed and tested for reliability and predictive capacity. The SSTTP is comprised of nine bipolar, research-based criteria (such as Friendly vs. Aloof); five deal with personality, two with logical thinking, and two with organization. The TCAP categorizes seven classroom activities, which together encompass four levels of behaviorally defined intellectual activity: managing, informing, thinking, and thinking about thinking. The instruments were subjected for three years to five treatment patterns (which varied as to number of observers and classroom visits per student teacher, time between observations, and observer training) to determine the effects of different observer patterns on reliability. Use of the TCAP revealed that only 10 percent of the classroom time was devoted to reasoning and thinking--which suggests a need for curriculum revisions. Relationships between SSTPP data and other data (such as grade point average) were not high enough to be of predictive value. The highest observer training; the next highest which approached acceptability (.80), was team observation. (Research which influenced the development of the instruments is reviewed and a 68-item bibliography is included.) (LP)

ED028999

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

BR-6-1321
PA-24
OE-BR

FINAL REPORT
Project No. 6-1321
Contract No. OEC-3-7-061321-0342

ISOLATING RELEVANT VARIABLES IN STUDENT TEACHER ASSESSMENT

Donald M. Sharpe
Division of Teaching
School of Education
Indiana State University
Terre Haute, Indiana 47809

January, 1969

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

SP002633

FOREWORD

A project of this scope and complexity is truly a team effort. The mobility of university personnel poses problems for an effort which extends over as many years as this one. Similarly, other professional responsibilities compete for attention and time. Unfortunately, Dr. Daniel Norton, who served as assistant director in designing the study, was forced to shift from a direct role to a consultative role because he left the university before the project was funded. Practical considerations required some modification of the original design. Professor Howard Rogers, a member of the Institute for Research in Human Behavior, assumed responsibility for part of the statistical analysis although he did not have the advantage of sharing in the original design and trial runs. Staff changes were also made among the field supervisors due to change in positions.

Valuable suggestions were made by Professors B. Othanel Smith, Harold Mitzel, Donald Medley, and Winslow R. Hatch and by Dr. Abraham Barnett, who also provided assistance in the interpretation of the factor analysis.

Particular appreciation is due Milton Firestone and the ISU Computer Center for many extra hours of work and computer time, to the ISU Audio-Visual Center for assistance in graphic presentation, to Jane Angell for editing, and to Carolyn Gurman for assistance in editing and for typing.

The staff is indebted to the many individuals in the public schools of Indiana and the more than two thousand student teachers who have at one time or another contributed to this project.

While the Director assumes responsibility for the final report the entire team must receive the credit for sharing in the identification of the problem, designing the instruments, and collecting and interpreting data. Howard Rogers was responsible for Chapter III and assisted with statistical analysis and interpretation.

TABLE OF CONTENTS

Page

Project Staff 7

Summary 9

Chapter

I Introduction. 11

II Background. 17

III Data Collection and Analyses. 39

IV The Teacher Classroom Activity Profile. 67

V Analysis of the Secondary Student Teacher Performance Profile 83

VI Conclusions and Implications. 114

Bibliography. 117

Appendices

A Guide for Assessment of Secondary Student Teacher Performance 123

B Instructions for Using the Teacher Classroom Activity Profile; Definitions and Illustrations of Terms. 137

C Anecdotal Reports Illustrating Criterion Behavior. 145

D Master Matrix of Correlation Coefficients 168

LIST OF TABLES

Page

1.1 Comparison of Grade-Point Averages of Study Groups with Those of Total Senior Class, Indiana State University 13

1.2 Distribution of Secondary Student Teachers and Students in the Study Group 1966-1968 by Subject Fields 14

2.1 Relationships Between the Schema Designed for This Study and Those Developed by Others for Organizing Ideas about Thinking and the Teaching Role 26

3.1 Treatment Classification 40

3.2 ANOVA Design for Treatments 1, 2, and 4 43

3.3 Formula for Derivation of Variance Components for Treatments 1, 2, and 4 44

3.4 Reliability of Seven Criteria Treatments 1, 2, and 4 44

3.5 Estimation of Variance Components in Treatment 1, 2, and 4 46

3.6 Product-Moment Correlations of Criterion Ratings in Treatments 1, 2, and 4 48

3.7 Projected Reliability of Three Treatments 49

3.8 ANOVA for Treatment 1 50

3.9 ANOVA for Treatment 2 51

3.10 ANOVA for Treatment 4 52

3.11 ANOVA Design for Treatment 3 54

3.12 Inter-Observer Reliabilities for Treatment 3 55

3.13 Product-Moment Correlations of Criterion Ratings in Treatment 3 55

3.14 Estimation of Variance Components in Treatment 3 56

3.15 Projected Study for Treatment 3 56

3.16 ANOVA for Treatment 3 57

3.17 Revised Instructions for Criteria VII and VIII 58

3.18 ANOVA Design for Treatment 5 59

3.19 Estimation of Variance Components 59

3.20 Reliabilities of Criteria for Treatment 5 60

3.21 Variance Component for Treatment 5 61

LIST OF TABLES (continued)

		Page
3.22	Anova for Changes Over Time	62
3.23	Factors of SSTPP	64
3.24	Communalities of Criteria	64
3.25	Orthogonal Factor Matrix (Varimax) of SSTPP	65
4.1	Mean Percentages of Time Devoted to Each Category of Teacher Classroom Activity--for Both Academic and Laboratory Classes	71
4.2	Mean Percentages of Time Devoted to Each Category of Teacher Classroom Activity for Laboratory-Type and Academic-Type Classes	72
4.3	Mean Percentages of Time Devoted to Each Category of Teacher Classroom Activity by Observers and by Type of Class	73
4.4	Analysis of Variance on Distribution of Classroom Time by Person Making Observation and by Type of Class	75
4.5	Mean Percentages of Time Devoted to the Various Categories of Teacher Classroom Activity by Subject Fields and Total for All Fields	78
4.6	Mean Percentages of Time Spent in Various Types of Classroom Activity During First and Second Observations	79
4.7	Proportion of Logical Entries Reported by Smith	80
4.8	Comparison of Percentages of Time Spent in Teacher Classroom Activity Categories and Hughes' Categories	81
5.1	Mean Estimates on Each Criterion of SSTPP Together with Standard Deviations by Observers	86
5.2	Analysis of Variance of Criterion Ratings by Semester and Platoon	93
5.3	Mean Estimates on Each Criterion of SSTPP Together with Standard Deviations by Subject Fields and Type of Class	96
5.4	Correlation Coefficients of Criterion Ratings on Secondary Student Teacher Personality Profile and Various Other Data on Student Teachers	99
5.5	Matrix of Correlation Coefficients Criterion IX (High Professional Potential) and Other Criteria	100
5.6	Correlation Coefficients of Criteria Ratings and Scores on the Sixteen Personality Factor Questionnaire	106
5.7	Correlation Coefficients of Criterion Ratings Made From Interviews and Observations	111

LIST OF FIGURES

Page

4.1 Cue Chart for Classifying Teacher Behavior on TCAP 70

4.2 Geometric Profile of Interaction Between Types of Observers and Types of Classes on Logical Thinking Category 76

5.1 Means, Standard Deviations, and Ranges of Criterion Estimates Made by Twelve Observers for Total Group and Experimental Group 84

5.2 Distribution of Estimates on Criteria by Platoons 89

5.3 Histogram of Raw Scores on Cornell Critical Thinking Test 102

5.4 Distribution of 16 P.F. Scores for Experimental Group, Non-Experimental Group, and College Undergraduates 105

PROJECT STAFF 1966-1969

Donald M. Sharpe, Project Director

Daniel P. Norton, Assistant Director
(Prior to October, 1966)
Consultant (1966-1969)

Howard Rogers, Research Assistant (1967-1969)

RESEARCH ASSOCIATES AND FIELD SUPERVISORS

Neil Amos
(1967-1968)

Wayne Beasley

Donald Bennett
(1967-1968)

Wenonah G. Brewer

Robert G. Dagley

Wilburn T. Elrod

Carroll D. Farrar

Marvin A. Henry

Samuel Mayhugh
(1967-1968)

Thomas A. Phillips

Chester S. Williams

IBM ANALYSIS AND TABULATION

Shirley Martin

Milton Firestone

SECRETARIAL ASSISTANTS

Carolyn W. Gurman
Helen K. Lebowitz

CONSULTANT

B. Othanel Smith

DEPARTMENTAL SUPERVISORS IN TCAP STUDY

Art	John Laska
Business	Wynnie Ford Ralph Mason
English	Laban Smith Mark Neville
Foreign Language	Frank Medley
Home Economics	Marjorie Jerry Anne Lee
Industrial Arts	Elmer Ciancone Raymond Keil Ethan Svendsen
Library Science	Norman Clark Nelle McCalla
Mathematics	Robert L. Kellems
Music	Wilburn Elrod
Physical Education--Men	Paul Wolf
Health and Safety	Walter Ney
Physical Education--Women	Louise Pound
Science:	
Biology	William Wert
Chemistry	William Kessel
General	Stanley Shimer
Geography	John Hook
Physics	John McCarthy
Social Studies	C. W. Engelland Jack Grantham
Special Education	R. B. Porter Margaret Rowe William E. Sherfey
Speech	Gladys Rohrig

SUMMARY

This investigation was concerned with developing and testing two instruments and their accompanying manuals for evaluating and reporting student teacher classroom behavior.

The "Secondary Student Teacher Performance Profile" consisted of the following bi-polar criteria suggested by contemporary research and educational thinking as being related to teaching success: I. Understanding, Friendly vs. Egocentric, Aloof; II. Planned, Organized, Responsible vs. Unplanned, Disorganized, Irresponsible; III. Stimulating, Imaginative, Surgent vs. Dull, Routine, Unimaginative; IV. Perceives Self as Competent vs. Perceives Self as Less Than Adequate; V. Has Mastery of Facts & Organizing Principles of Field vs. Has Only Minimum Knowledge of Field; VI. Communicates Well & Empathetically vs. Communicates Ineffectively & Perfunctorily; VII. Classroom Discourse Characterized by Reasoning or Creative Thinking vs. Classroom Discourse Characterized by Simple Recall; VIII. Directs Attention to Process of Thinking vs. Fails to Call Attention to Process of Thinking; IX. High Professional Potential vs. Low Professional Potential.

The "Teacher Classroom Activity Profile" provided the following seven categories of activity among which the student teacher's time was distributed: MN--Management-Non-Learning; ML--Management-Learning; P--Presentation; R--Recitation; D--Discussion/Random; LT--Logical Thinking; and TP--Thinking Process.

Student teachers were evaluated on the Secondary Student Teaching Performance Profile criteria under five experimental "treatments" to determine the reliability of the ratings. The intra-observer reliabilities were low. The intra-observer treatments revealed some differences but none of them produced a high level of reliability. The inter-observer reliabilities with five observers on a team approached an acceptable level. Where a special training procedure was used the reliabilities were high.

Some relationships were found between criterion ratings and grade point averages, estimated grade point averages, and supervising teachers' ratings. No correlations were found between criterion ratings and scores on the Scholastic Aptitude Test, and the Cornell Critical Thinking Test. Some scores on the 16 Personality Factor Questionnaire were found to be statistically, significantly different from zero. They were so low as to have little predictive value; however, they do suggest that further study might be fruitful.

As a result of cooperative study and in-service training, the staff members felt that they were able to make judgments of sufficient reliability as to be helpful in modifying the student teacher's professional education and influencing his job placement. Criteria VII and VIII concerned with reasoning and thinking proved the most difficult, but at the

same time continued to present a challenge. The fact that only ten percent of classroom time was devoted to reasoning and higher level intellectual activities suggested the need for curriculum revision in the secondary schools and in the teacher preparatory institutions.

While further refinements of the techniques are necessary there is no evidence that evaluative instruments and techniques used in this study are any less reliable than current practices of credit counting, giving grades and making generalized evaluations.

The Secondary Student Teacher Performance Profile and the Teacher Classroom Activity Profile give promise of being effective aids in the supervision and evaluation of student teachers.

CHAPTER I

INTRODUCTION

The Problem

This study was an attempt to determine the extent to which it is possible to evaluate the performance of selected secondary student teachers on criteria drawn from contemporary research on teacher behavior in the typical student teaching classroom and, further, to discover how such evaluations relate to other data descriptive of the student teacher.

The Setting

The Program

Secondary Student Teaching at Indiana State University is a full-time, off-campus, eight-week experience. One half of the professional semester, which occurs during the student's senior year, is spent in the field and the other half on campus. Platoon I students are in the field the first eight weeks of the year and on campus the second eight weeks. Platoon II students are on campus the first eight weeks and in the field the second. Similar arrangements are made for Platoons III and IV during the second semester.

Prerequisites to the professional semester are General Psychology, Psychology of Childhood and Adolescence, and Teaching and Learning in the Secondary School. Some students take the methods course in their subject field before the professional semester. During one-half of the professional semester students enroll in supervised teaching for six hours of credit, one of which covers the orientation and evaluation seminars. The other half-semester he enrolls in campus courses--Educational Technology (1 hour credit), Problems in Secondary Teaching (2 hours), Foundations of Secondary Education (2 hours), and a methods course in the subject field.

In the regular program at Indiana State University student teachers are visited at least three times by the college representative. The first visit is made early in the assignment before students are actively engaged in teaching, and is primarily concerned with administrative matters and personal adjustment to the situation. On the second and third visits the supervisor observes the student as he teaches.

The Staff

Administration and supervision of the student teaching is provided by full-time staff members of the Division of Teaching and shared by Departmental Supervisors who teach the special methods courses. At the present time the Division of Teaching consists of eleven full-time staff members and twenty-one part-time Departmental Supervisors. All members of the Division of Teaching staff have academic rank--one Instructor, four Assistant Professors, three Associate Professors, and three Professors. During the academic year they devote their full time to student teaching, each working with 17 to 20 students each of the four platoons, in one section of the state of Indiana. Staff meetings are held each Monday on campus. The staff is also on campus for the orientation seminars, application conferences, and evaluation week. During the summer term most members of the staff teach education courses.

Departmental Supervisors, who are members of the subject-matter departments, teach the special methods courses, work with the Division of Teaching staff in placing student teachers, and actually supervise their students who are assigned in the local community. Approximately one-fifth of the student teachers are supervised by Departmental Supervisors who are specialists in the subject field and four-fifths are supervised by the Division of Teaching staff who are generalists or, rather, specialists in Pedagogy.

The members of the Division of Teaching staff participated in all phases of this study. The Departmental Supervisors participated only in the phase concerned with the analysis of classroom activity, reported in Chapter V.

The Student Population

All student teachers involved in this project were randomly selected. Tables 1.1 and 1.2 provide selected data on these students and on other members of the senior class for the two years 1966-1967 and 1967-1968.

As shown in Table 1.1, the population involved in the study had a slightly higher proportion of females than that of the senior class. The grade-point percentiles for the study group were slightly higher than those of male seniors and slightly lower than those of female seniors. The differences with respect to sex and grade-point average between the student teachers involved in the study and other members of the senior class were relatively slight.

Although no attempt was made to provide proportionate representation by subject-teaching-field, Table 1.2 shows that no serious imbalance occurred with respect to subject fields. Approximately fifty percent of the student teachers enrolled each semester were included in the study.

TABLE 1.1

COMPARISON OF GRADE-POINT AVERAGES OF STUDY GROUPS
WITH THOSE OF TOTAL SENIOR CLASS, INDIANA STATE UNIVERSITY

Grade-Point Average	Seniors, ISU *				Student Teachers in Research Study				
	1966-67		1967-68		Treatments				
	Male	Female	Male	Female	1	2	3	4	5
Mean--M					2.66	2.90	2.73	2.60	2.62
Percentiles									
75	2.82	3.12	2.85	3.12	2.97				
Median--50	2.48	2.74	2.50	2.74	2.58				
25	2.23	2.41	2.22	2.38	2.28				
Number N =	827	656	1310	1089	538	20	42	23	5
Percentage Male	55		55		47	45	52	44	60

* Data supplied by Indiana State University Office of Institutional Research and Testing. GPA mean scores were not available for all seniors.

TABLE 1.2

DISTRIBUTION OF SECONDARY STUDENT TEACHERS AND STUDENTS IN THE STUDY GROUP INDIANA STATE UNIVERSITY 1966-1968 BY SUBJECT FIELDS

	Students not in Study		Students in Study	
Department	Supervised by Department	Supervised by Division of Teaching	Supervised by Division of Teaching	Total
Art	10	13	16	39
Business	15	13	49	77
English	52	16	78	146
For. Lang.	15	8	37	60
Home Ec.	13	7	34	54
Ind. Ed.	59	44	53	156
Math	29	13	40	82
Music	8	30	42	80
Phys. Ed.	56	43	80	179
Science	28	10	35	73
Soc. St.	71	20	72	163
Sp. Ed.	40	26	28	94
Totals	396	243	564	1203

The Observational Patterns

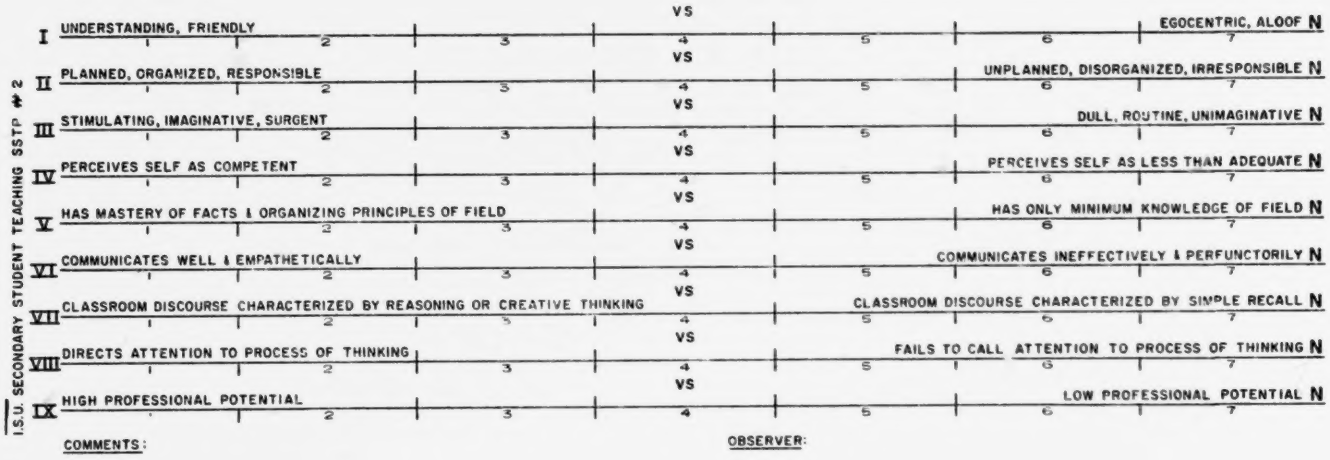
Four patterns of observing student teacher performance were planned in the experimental design. These patterns, which will be called Treatments, were as follows:

> Treatment 1: A member of the Division of Teaching staff observed a student teacher for a full class period during the third or fourth week of the student's field experience and again during the seventh or eighth week. This pattern was similar to the regular university student teaching program and constituted a quasi-control group.
>
> Treatment 2: A staff member observed a student teacher four times--once each in the second, fourth, sixth, and eighth weeks.
>
> Treatment 3: A team of two staff members observed a student in a single class during the fifth or sixth week.
>
> Treatment 4: A staff member observed a student in two different classes on the same day.
>
> Treatment 5: A replication of Treatment 3 during the summer of 1968.

The Instruments

The basic instrument, Secondary Student Teacher Performance Profile (SSTPP), consists of nine bi-polar criteria on which the observer rates the student's performance on a seven-point scale. A copy follows.

Secondary Student Teacher Performance Profile



I.S.U. SECONDARY STUDENT TEACHING SSTP #2

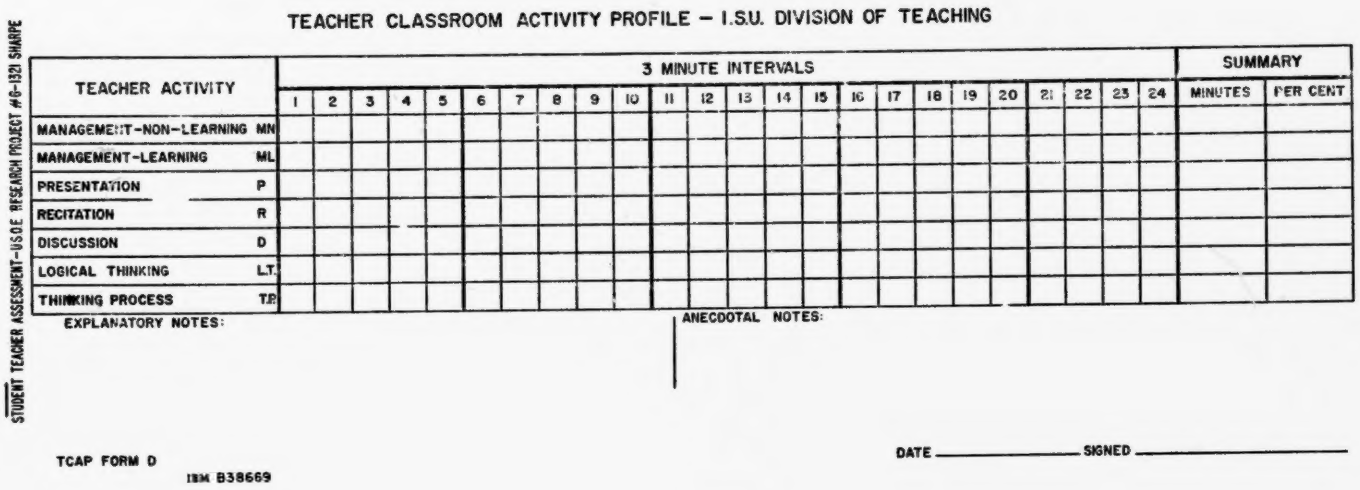
I	UNDERSTANDING, FRIENDLY	1	2	3	VS 4	5	6	7	EGOCENTRIC, ALOOF	N
II	PLANNED, ORGANIZED, RESPONSIBLE	1	2	3	VS 4	5	6	7	UNPLANNED, DISORGANIZED, IRRESPONSIBLE	N
III	STIMULATING, IMAGINATIVE, SURGENT	1	2	3	VS 4	5	6	7	DULL, ROUTINE, UNIMAGINATIVE	N
IV	PERCEIVES SELF AS COMPETENT	1	2	3	VS 4	5	6	7	PERCEIVES SELF AS LESS THAN ADEQUATE	N
V	HAS MASTERY OF FACTS & ORGANIZING PRINCIPLES OF FIELD	1	2	3	VS 4	5	6	7	HAS ONLY MINIMUM KNOWLEDGE OF FIELD	N
VI	COMMUNICATES WELL & EMPATHETICALLY	1	2	3	VS 4	5	6	7	COMMUNICATES INEFFECTIVELY & PERFUNCTORILY	N
VII	CLASSROOM DISCOURSE CHARACTERIZED BY REASONING OR CREATIVE THINKING	1	2	3	VS 4	5	6	7	CLASSROOM DISCOURSE CHARACTERIZED BY SIMPLE RECALL	N
VIII	DIRECTS ATTENTION TO PROCESS OF THINKING	1	2	3	VS 4	5	6	7	FAILS TO CALL ATTENTION TO PROCESS OF THINKING	N
IX	HIGH PROFESSIONAL POTENTIAL	1	2	3	VS 4	5	6	7	LOW PROFESSIONAL POTENTIAL	N

COMMENTS:

OBSERVER:

A supplementary instrument, Teacher Classroom Activity Profile (TCAP), is an instrument for recording the way the teacher spends his time in the classroom. Seven discrete categories of behavior are identified. The observer records a continuous line from the beginning until the end of the observational period using one minute intervals. A copy follows.

Teacher Classroom Activity Profile



STUDENT TEACHER ASSESSMENT-USOE RESEARCH PROJECT #6-1321 SHARPE

TEACHER CLASSROOM ACTIVITY PROFILE – I.S.U. DIVISION OF TEACHING

TEACHER ACTIVITY	3 MINUTE INTERVALS																								SUMMARY	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	MINUTES	PER CENT
MANAGEMENT-NON-LEARNING MN																										
MANAGEMENT-LEARNING ML																										
PRESENTATION P																										
RECITATION R																										
DISCUSSION D																										
LOGICAL THINKING L.T.																										
THINKING PROCESS T.P.																										

EXPLANATORY NOTES:

ANECDOTAL NOTES:

TCAP FORM D

IBM B38669

DATE ____________ SIGNED ____________

Treatment of Data

The data were first analyzed to ascertain the effects of different observational patterns on reliability. Second, a factor analysis was employed to determine the extent to which common elements were present over the nine criteria. Finally, the ratings on the nine criteria were examined to determine if they were related to the following data: grade-point average, estimated grade-point average, estimates made from an interview, evaluative reports by cooperating teacher, SAT score, score on the Cornell Critical Thinking Test, scores on the 16 Personality Factor Questionnaire, and time distribution on TCAP.

Schedule

The study covers a five-year period. Beginning in 1964, two years were spent studying the research, agreeing upon evaluative criteria, designing the instruments, training observers, and running a pilot study. The next two years were devoted to collecting data. The final year was spent in analyzing the data and writing the report.

CHAPTER II

BACKGROUND

While this project has its roots in the cumulative experiences of all those who participated in it, it is directly related to an earlier attempt to improve the evaluation of student teachers, which was abandoned. In 1950 the following items were added to the Indiana State College Student Teacher Evaluation Form: "Comment on the student's professional philosophy, attitudes, and understandings: Does he understand the crucial role schools play in maintaining and extending the democratic way of life? Does he try to develop those values, habits and skills necessary for our democratic society? Does he use the scientific method?" Those who completed the form resisted responding to these items so vigorously that the items were revised to read: "Comments on the Student's Understanding, Attitudes, Philosophy, etc." No one denied the relevance of the questions but almost everyone felt inadequate to the challenge of answering. In a real sense this project is a second attempt to improve the evaluation of student teacher performance by developing some new instruments and procedures.

Although previous attempts to quantify and appraise teacher effectiveness have been only partially successful, the scientific upgrading of education requires nothing less.

The rational development of a professional program for teacher preparation requires that professional competencies be identified and ways of assessing the degree of competence be developed. Conant suggests that the student teacher's performance in the classroom should be the chief factor in granting him a certificate to teach. Hopefully, teacher certification could come to be based on professional competence rather than on tabulation of the number of courses taken.[1]

The professional role of the college supervisor in visiting a student teacher's classroom has not been adequately described. Too often he concentrates his attention on the psychological state of the student rather than on his actual teaching performance. Too much supervisory attention is devoted to praise or criticism and too little to analysis and interpretation.

---

[1] J. B. Conant, The Education of American Teachers (New York: McGraw-Hill, 1963).

Current practices in the evaluation of student teaching reflect the uncertainty about purposes and methods of teacher preparation which has characterized the profession for the past half century. Most instruments for evaluating student teaching concentrate on generalized qualities of personality and almost ignore pedagogical competence. Recent attempts to define the act of teaching and recent research on teaching skills and teacher competencies suggest the possibility of determining criteria upon which a teacher's performance can be evaluated. The student teaching experience provides an opportunity to apply the evaluative criteria during the student's pre-service education.

The staff hoped to develop instruments and techniques for evaluating student teacher performance which could be used in a typical student teaching situation. At the same time it was proposed to study experimentally the relationships which obtain among the criteria when they are employed in a variety of observational situations and sequences to establish the reliability and operational utility of ratings so made by college supervisors.

Research on teacher behavior suggests that certain criteria on an acturial basis are related to teaching success. This study attempted to apply such criteria in actual classroom situations to individual student teachers.

Most research dealing with teacher classroom behavior involves such complex instrumentation and requires such detailed recording, coding, and analysis that only full-time researchers or specialists in observation can conduct it. The staff hoped to develop a technique which could be employed by specially trained college supervisors as they worked in a student teaching program. Research and curriculum improvement were to be "double yoked."

Rationale

This study rested upon ~~four~~ implicit assumptions.

First, since teaching is behavior, it can be analyzed and changed. It was recognized that teaching behavior is complex and difficult to analyze but it must be so analyzed if teaching behavior is to be improved.

Second, a study of the behavior of teachers should include a focus upon intellectual operations as a necessary supplement to the many recent research projects which focus primarily upon classroom climate and interaction. While teaching is interaction between teacher and pupil, this study will be concerned primarily with the teacher's behavior and only incidentally with the pupils' reactions.

Third, relationships between effective teaching and certain characteristics and behaviors which have been identified on an actuarial basis could be applied on a clinical or individual basis.

Fourth, evaluation of teacher classroom behavior can be made by trained observers without employing the time-consuming collecting, coding, and analyzing of detailed protocol material by full-time researchers.

It was recognized that each of the assumptions posed problems and presented both theoretical and practical difficulties. The staff recognized that participation in such a project would likely result in improved supervisory practice and curriculum reorganization. They recognized that "very little is known for certain about the nature and measurement of personality, or about the relation between teacher personality and teaching effectiveness,"[2] but they saw much merit in pursuing some of the implications of recent research. The importance of the emotional climate in the classroom was not minimized but the staff felt that the intellectual aspects of teacher behavior were of greater significance. They felt, in the words of John Fischer, that before teaching can attain full professional status it will be "necessary first to specify the knowledge and skill that distinguish a professional teacher from other well-intentioned people who get along with students."[3]

The problem was to eclectically identify criteria for evaluating teacher behavior as reported in recent research and then develop a technique for using these criteria in assessing student teacher behavior.

The nine criteria agreed upon for this study were assumed to be valid on the basis of prior research and/or expert opinion. No attempt was made to evaluate the effect the teacher had upon pupil learning.

In summary, the premises of this study were that a highly trained college supervisor can (a) improve his helpfulness to the student teacher, (b) provide more accurate prognostic assessments, and (c) give better direction to the institution in its program development if he employs criteria which are drawn from research and which are behaviorally defined rather than if he uses the vague, non-specific, and poorly-defined criteria which are usually employed in student teaching evaluation.

---

[2] N. L. Gage (ed.), Handbook of Research on Teaching (American Educational Research Association, Chicago: Rand McNally, 1963), p.574.

[3] John H. Fischer, review of Turmoil in Teaching by T. M. Stinnett, Phi Delta Kappan, Vol. 50, No. 2 (October, 1968), p. 119.

Defining "Teaching"

In the absence of an agreed-upon theory of teaching, the staff accepted the paradigm developed by B. O. Smith which describes the teaching cycle as the sequential process of interacting between teacher and learner which consists of the teacher's perceiving, diagnosing, and reacting to the pupil who in turn perceives, diagnoses, and reacts to the teacher.

$$P_t \rightarrow D_t \rightarrow R_t \rightarrow P_p \rightarrow D_p \rightarrow R_p \rightarrow P_t \rightarrow D_t \rightarrow \text{etc.}$$

Where P = Perceiving, D = Diagnosing, R = Reacting and the
Subscripts t = teacher, p = pupil[4]

It is obvious that such a paradigm rejects the concept of teaching as merely the imparting of knowledge or the mastery of a particular method. This paradigm suggests that special preparation to teach is necessary in order to help the teacher perceive and diagnose accurately and make the most promising reaction. It recognizes the complexity of the teaching act and the infinite varieties of styles and strategies which can be employed, but focuses upon meaningful, finite agglomerates.

The staff in surveying research on teaching was well aware that little is known for certain about the nature and measurement of teacher personality and teaching effectiveness. However, research and expert opinion suggested the nine criteria which they were content to apply to teacher performance in the classroom.

Defining "Thinking"

While examining the intellectual aspects of teacher behavior constituted the chief raison d'être it also posed the most serious problems. The problem of defining thinking in behavioral terms has not yet been solved. In common use thinking has meant anything that goes on in the head from day-dreaming to designing a formula for inter-stellar travel. It was necessary to develop a functional concept of thinking in the classroom.

This section attempts to share some of the concepts related to thinking as the staff agreed upon a vocabulary and a rubric necessary for effective communication. This is not presumed to be a definitive analysis of thinking.

George Bernard Shaw is reported to have said that there are thousands of ways to write a poem and every one of them is right. Similarly, there are hundreds of ways of thinking about thinking and every one of them has its justification. It is pointless to argue the merits of one system of categorizing thinking over another system; however, it was necessary to agree upon the categories and terms to be used in this project.

---

[4] B. O. Smith, "A Concept of Teaching," Chapter Six, Language and Concepts in Education by B. O. Smith and R. H. Ennis (Chicago: Rand McNally, 1961).

Webster's Unabridged Dictionary includes the following definitions of the verb "to think": (1) to form or have in the mind; conceive: (2) to hold in one's opinion; judge; consider: (3) to believe; surmise; expect: (4) to determine, resolve, work out, etc. by reasoning: (5) to have an intent; purpose; intend: . . . .

While the development of thinking has always been accepted as an important educational objective, its implementation in classrooms has been haphazard. An examination of curricula and classroom practices leads to the conclusion that thinking is frequently only a by-product of studying certain subjects or accumulating information. Little attention has been paid to helping the teacher learn to assist students in improving their thinking. The examination of the literature in the field revealed no adequate definitions or models for dealing with thinking but did provide help in developing a workable system.

A very simple system of categorization which is widely used identifies two types of thinking as (1) inductive or (2) deductive. Recently Gallagher and others emphasize convergent vs. divergent thinking.

The Scientific Method, or the Method of Inquiry, is used to describe the processes of (1) formulating an hypothesis, (2) collecting data, (3) checking the hypothesis against the data, and (4) drawing a conclusion.

A similar system is commonly referred to as Problem Solving and usually refers to the processes involved in (1) identifying a problem, (2) hypothesizing a solution, (3) collecting data, (4) checking the data against the hypothesis--revising the hypothesis in light of the implications of the data, and (5) applying the results to the solution of the problem. These are closely related to Dewey's steps in the thinking process which have been identified as (1) sensing a problem, (2) identification of the nature of the problem, (3) searching for possible solutions, (4) analyzing the adequacy of tentative solutions, and (5) testing the most promising of the solutions by action.

The Progressive Education Association's Eight-Year Study focused on the development of critical thinking. Three objectives of thinking were identified: (1) the ability to infer generalizations from specific data, (2) the ability to apply known principles in explaining new situations or predicting consequences, and (3) the ability to do critical thinking *per se*.[5]

---

[5] E. R. Smith and R. W. Tyler, *Appraising and Recording Student Progress* (New York: Harper, 1942).

B. O. Smith in A Study of the Logic of Teaching identified the following twelve categories of logical operations in teaching: (1) Defining, (2) Describing, (3) Designating, (4) Stating, (5) Reporting, (6) Substituting, (7) Evaluating, (8) Opining, (9) Classifying, (10) Comparing and Contrasting, (11) Conditional Inferring, and (12) Explaining. He further subdivided Categories 1, 3, and 12 and recognized that some attention is devoted to managing the classroom.[6]

John C. Wilson suggested that prose statements can be classified into the following types: (1) Imperatives and Attitude Statements, (2) Empirical Statements, (3) Analytical Statements, (4) Value Statements, and (5) Metaphysical Statements.[7]

Bloom, in his work in developing a system for classifying educational objectives, identified the following categories within the cognitive domain: (1) Memory (Knowledge), (2) Translation, (3) Interpretation, (4) Application, (5) Analysis, (6) Synthesis, (7) Evaluation.[8]

Mary Jane Aschner suggested that one way to approach the problem of logic and language in the classroom would be to classify the thinking under the following rubric: (1) Remembering, (2) Reasoning, (3) Evaluating or Judging, and (4) Creative Thinking.[9]

Donald Medley classified teacher statements in coding behavior with OScAR 4V into six major categories with three areas: (1) Affective Statements, (1.1) Considering, or (1.2) Rebuking; (2) Substantive Statements, (2.1) Informing, or (2.2) Problem Structuring; and (3) Procedural Statements, (3.1) Directing, or (3.2) Describing.[10]

---

[6] B. O. Smith, Mary Jane Aschner, and Milton Meux, A Study of the Logic of Teaching, A report of the first phase of a five years' research project, The Logical Structure of Teaching and the Development of Critical Thinking, Project #258 (7257), U.S. Office of Education (Urbana, Ill.: Bureau of Educational Research, College of Education, University of Illinois). (Dittoed.)

[7] John Wilson, Language and the Pursuit of Truth (Cambridge, England: University of Cambridge Press, 1960).

[8] Benjamin Bloom (ed.), Taxonomy of Educational Objectives (New York: Longmans, Green, 1958).

[9] Mary Jane Aschner, "The Language of Teaching," Teachers College Record, Vol. 61 (February, 1960), pp. 251-52.

[10] Donald M. Medley, "Coding Behavior with OScAR 4V," pre-publication draft (Princeton, N.J.: Educational Testing Service). (Mimeographed.)

Louis E. Raths and his colleagues suggested the following types of thinking operations: (1) Comparing, (2) Summarizing, (3) Observing, (4) Classifying, (5) Interpreting, (6) Criticizing, (7) Looking for Assumptions, (8) Imagining, (9) Collecting and Organizing Data, (10) Hypothesizing, (11) Applying Facts and Principles, (12) Decision-Making, and (13) Designing Projects or Investigation.[11]

Gallagher, Aschner, and Jenne[12] based their classification upon Guilford's[13] model of intellectual performance to arrive at (1) Cognitive Memory, (2) Convergent Thinking, (3) Divergent Thinking, and (4) Evaluative Thinking. To these four they added (5) Routine.

Woodruff developed a schema showing the relationships between mental processes in learning and their related processes in teaching. Mental processes are classified as (1) Level A--Perception, Stimulation; (2) Level B--Inductive; (3) Level C--Inductive Tryout; and (4) Level D--Deductive--Analysis and Creation. The corresponding processes of teaching are: (1) Level A--Show; (2) Level B--Discuss; (3) Level C--Apply; and (4) Level D--Guide Original Thinking.[14]

The Ford Foundation experimental intern program at Stanford University studied the problem of identifying criteria for appraising intern trainees.[15] Unfortunately their work did not come to the attention of the Indiana State University staff until too late to influence the design or operation of this study.

---

[11] Louis E. Raths, Selma Wassermann, Arthur Jonas, and Arnold M. Rothstein, Teaching for Thinking--Theory and Application (Columbus, Ohio: Charles E. Merrill Books, Inc., 1967).

[12] James J. Gallagher, Mary Jane Aschner, and William Jenne, Productive Thinking of Gifted Children in Classroom Interaction, CEC Research Monograph Series B., No. B-5 (Washington, D.C.: The Council for Exceptional Children, NEA, 1967).

[13] J. P. Guilford, "Three Faces of Intellect," American Psychologist, Vol. 14 (1959), pp. 469-79.

[14] Asahel D. Woodruff, Basic Concepts of Teaching (San Francisco: Chandler, 1961), pp. 126-27.

[15] Robert N. Bush, Norman J. Boyan, and Dwight W. Allen, Stanford University School of Education Secondary Teacher Education Program, 1965-1966. Resident Supervisors Handbook (Palo Alto). (Offset.)

The Stanford Teacher Competence Appraisal Guide 1964-65 rated interns on a seven-point scale with respect to thirteen teacher roles clustered under four major headings: Aims--Clarity of Aims, Appropriateness of Aims; Planning--Organization of the Lesson, Selection of Content, Selection of Materials; Performance--Beginning the Lesson, Clarity of Presentation, Pacing of the Lesson, Pupil Participation and Attention, Ending the Lesson, Teacher-Pupil Rapport; Evaluation--Variety of Evaluative Procedures, Use of Evaluation to Improve Teaching and Learning.

While both the Stanford program and the Indiana State University program appear to have grown out of similar rationales and to use a similar approach, they differ markedly in the criteria selected and the instruments developed. The Stanford Teacher Competence Appraisal Guide as contrasted with the Indiana State University Secondary Student Teacher Performance Profile calls for judgments on thirteen elements as contrasted with nine. The Stanford approach breaks down teacher behavior into more specific actions. The Indiana State University SSTPP form is restricted to evaluating the student on the basis of what he does in the classroom whereas the Stanford Guide includes ratings of aims and planning. It is possible that the interest at Stanford in identifying the micro-aspects of teaching shaped the structuring of the Guide. The criteria appear to be related to the Stanford conception of the various discrete roles a teacher plays. The Indiana State criteria, by contrast, grew out of the staff's perceptions of the kinds of behavior research has suggested are related to teaching success.

Hilda Taba presented an excellent analysis of the problems inherent in studying the thinking process and teaching strategies.[16] Unfortunately, her study was not available to the staff until most of the work had been completed.

In the absence of any commonly agreed upon system for organizing ideas about intellectual activity the staff developed its own. The Indiana State University system recognized that a great variety of intellectual activities are involved in teaching and that it would be necessary to develop some operational definitions which would permit communication among all concerned. No claim for its superiority over other systems is made. It is not presented as a definitive taxonomy of the thinking process.

Wittgenstein's view that "the meaning of a word is its use in the language"[17] was accepted.

---

[16] H. Taba, S. Levine, and F. F. Elzey, Thinking in Elementary School Children, USOE Cooperative Research Project No. 1574 (San Francisco State College, 1964).

[17] Ludwig Wittgenstein, Philosophical Investigations (Oxford: Basil Blackwell, 1958), p. 232.

The model which evolved from the staff study identified four levels of intellectual activity which could be categorized from observations of teacher behavior.

Levels of Intellectual Activity	Teacher Activity
Level One: Managing	Directing, Commanding, Managing, Rebuking, Accepting Students. These are usually imperative statements.
Level Two: Informing	Presenting Information, Reporting, Describing, Recalling from Memory. Getting students to do these. These are usually empirical statements.
Level Three: Thinking	Summarizing, Classifying, Explaining, Reasoning and Evaluating. Getting students to do these. These are usually relational statements.
Level Four: Thinking Process	Thinking about Thinking; Analyzing the Processes of Thinking Going on in Others; Recognizing the Roles Played by Assumptions, Biases, and Values; Locating Sources of Misunderstanding. Getting students to do these. These are usually logical statements.

This analysis of thinking proved helpful in identifying the various categories for classifying student teacher behavior on the Teacher Classroom Activity Profile. It also provided the basis of making evaluative ratings on Criteria VII and VIII of the Secondary Student Teacher Performance Profile.

The schema which follows suggests relationships between this system for looking at the teacher's intellectual activities and the staff's interpretation of the way some other scholars have organized their thinking. The relationships are approximations and tangential. The purpose of showing them here is to give added meaning to this system of classification rather than to compare one with another.

TABLE 2.1

RELATIONSHIPS BETWEEN THE SCHEMA DESIGNED FOR THIS STUDY AND THOSE DEVELOPED BY OTHERS FOR ORGANIZING IDEAS ABOUT THINKING AND THE TEACHING ROLE

Sharpe and Staff	Wilson[7]	Gallagher[12]	Raths[11]	Aschner[9]
Level One: Managing	Imperative and Attitude Statements	Routine		
Level Two: Informing	Empirical Statements	Cognitive Memory	Observing Collecting Data	Remembering
Level Three: Thinking	Empirical Statements	Convergent Thinking Divergent Thinking Evaluative Thinking	Summarizing Comparing Classifying Imagining Hypothesizing Applying Facts and Principles Making Decisions Designing Projects	Reasoning Evaluating or Judging Creative Thinking
Level Four: Thinking Process	Analytical Statements		Looking for Assumptions	

TABLE 2.1

RELATIONSHIPS BETWEEN THE SCHEMA DESIGNED FOR THIS STUDY AND THOSE DEVELOPED BY OTHERS FOR ORGANIZING IDEAS ABOUT THINKING AND THE TEACHING ROLE
(Continued)

Sharpe and Staff	Smith[6]	Medley[10]	Bloom[8]	Woodruff[14]
Level One: Managing	Directing and Managing	Affective Statements (Considering and Rebuking) Procedural Statements Directing and Describing		
Level Two: Informing	Describing Stating Reporting	Substantive Statements--Informing	Memory	Level A--Show
Level Three: Thinking	Defining Designating Substituting Evaluating Opining Classifying Comparing and Contrasting Conditional Inferring Explaining	Substantive Statements--Problem Structuring	Translation Interpretation Application Analysis Synthesis	Level B--Discuss Level C--Apply Level D--Guide Original Thinking
Level Four: Thinking Process			Evaluation	Some Aspects of Guiding Original Thinking

Developing the Instruments

The basic instruments,"A Guide for Assessment of Secondary Student Teacher Performance (Tentative)"* and the "Secondary Student Teacher Performance Profile" (SSTPP) were cooperatively developed by the staff as an outgrowth of a study begun in 1963.

The staff studied the literature and research on teacher competence and teacher behavior in the classroom. Three of the researchers, Smith, Mitzel, and Flanders, visited the campus and consulted with the staff. During fortnightly staff meetings the techniques for reporting judgments and the procedures for making observations were agreed upon. During the spring semester of 1964 and both semesters of 1964-65 the staff used the instrument on its regular supervising visits. Changes were made in Criteria IV and VIII and the Guide was refined and amplified. Further additions and modifications were made on the Guide during the project.

The Guide describes the instrument, Secondary Student Teacher Performance Profile (SSTPP), and how it is to be used by college supervisors in reporting observations of student teachers. It identifies the theoretical and research foundations for each of the nine criteria of the SSTPP. It provides operational definitions of terms by giving examples of behavior which serve as clues for rating student teachers' performance on the various dimensions and criteria of SSTPP. It describes the procedures to be followed in making observations.

The Secondary Student Teacher Performance Profile provides nine criteria upon which the supervisor will rate the student on a seven-point continuum. The classroom behavior of teachers provides the organizing principle for this instrument. The nine criteria were drawn from various research studies which had been designed with differing logical premises. The criteria, therefore, are eclectically rather than logically derived.

A second instrument, The Teacher Classroom Activity Profile (TCAP), together with a guide for its use and a glossary of terms, was cooperatively developed by the staff to provide a means for recording the ways in which the teacher spends classroom time.**

---

* A copy of the Guide is included as Appendix A.

** The development of TCAP is further explained in Chapter IV. A copy of the instrument, the guide, and glossary of terms appear as Appendix B.

Review of Research

Research related to this study falls into two classes. The first covers observation of teacher behavior; the second is concerned with criteria which might be employed during observation. These classes are not discrete. Some research has significance in both classifications.

Research Related to Observing Teacher Behavior

The work of B. Othanel Smith in developing the concept of teaching as "a system of actions intended to induce learning"[18] and his study of the logic of teaching[19] provided the initial stimulus for this study. Conant's report, which might more properly be called his proposals rather than research, suggesting that certification be based on a person's ability to teach, has highlighted the urgency of discovering a means whereby the teacher's ability to teach can be objectively determined in the classroom.[20]

Smith found that it was possible to classify teacher behavior in the classroom into three categories: Expressive Behavior, Performative Behavior, and Linguistic Behavior. He was also able to classify logical operations that went on in the classroom by careful transcription of classroom discourse and subsequent classification of what he called "episodes."

The work of the NEA Commission on Teacher Education and Professional Standards as reported by Margaret Lindsey played an important role by identifying the problem of teacher professional competence and challenged the staff to embark on this project.[21]

Two publications were most influential in designing the procedures employed in this investigation. They are the chapter contributed by Medley and Mitzel[22] to the Handbook of Research on Teaching and the report of an investigation conducted by Wilk and Edson.[23] In addition to serving as models for the performance of this investigation, they provided technical guidance for performing the research which was proposed.

---

[18] B. Othanel Smith and Robert H. Ennis, Language and Concepts in Education (Chicago: Rand McNally, 1961).

[19] Smith, Aschner, and Meux, op. cit.

[20] Conant, op. cit.

[21] Margaret Lindsey, New Horizons for the Teaching Profession (Washinton, D.C.: National Commission on Teacher Education, NEA, 1961).

[22] Donald M. Medley and H. E. Mitzel, "Measuring Classroom Behavior by Systematic Observation," Chapter 6, in N. L. Gage, op. cit.

Medley and Mitzel stated, "Certainly there is no more obvious approach to research on teaching than direct observation of the behavior of teachers while they teach and pupils while they learn . . . . That observational data are difficult and expensive to obtain is a fact that cannot be denied."

Writing in 1958 in the Journal of Experimental Education, they discussed the relative merits of analysis of variance and correlation analysis as techniques for estimating reliability coefficients and concluded that analysis of variance should be adopted as the preferred technique. This conclusion was restated and expanded in their handbook chapter. In the latter, it was suggested that mixed statistical models were appropriate for observational studies of teacher behavior and that observers' tasks should be kept as nearly objective as possible. The studies reviewed by Medley and Mitzel incorporated criteria employed through the 1950's. The broadly generalizable model which they presented provided for analysis of variance with items as factors in the analysis.

The investigation reported by Wilk and Edson included many details which are relevant to this investigation. The subjects were 36 elementary student teachers. They were observed by regular supervisors of student teaching and by specially trained observers. The OScAR III and the Minnesota System of Interaction Analysis were employed as student teacher behavior criterion measures. Pertinence of the study to this study may be recognized in a question which the investigation attempted to answer. It was:

> Using certain teaching behaviors as the basis for comparison, what is the relationship of ratings of student teachers by student teaching supervisors using their established observational methods and the ratings made by research observers using the methods designed for this study?

The Wilk and Edson study did not employ a mixed model for analysis. However, their model was close to the mixed model in function. It seems likely that results of their analysis can provide direction for persons interested in research on observational problems in student teacher assessment. Among the conclusions arrived at from their analyses were the following:

> On the basis of the admissions interview and data available at the time of admission, counselors could not predict successfully which students would exhibit integrative classroom behavior as student teachers.
>
> The student teaching supervisor and research observers . . . exhibited a high degree of general agreement when rating

---

[23] Roger E. Wilk and William H. Edson, A Study of the Relationships Between Observed Classroom Behaviors of Elementary Student Teachers: Predictors of Those Behaviors, and Ratings by Supervisors (Minneapolis: College of Education, University of Minnesota, January, 1962).

> either personality variables or judging classroom performance. However, there is little relationship between the scores assigned to the need dominance factor.
>
> To the extent that there is lack of agreement between the ratings of supervisors and observers one may conclude that this is due to the unaccounted-for variance, i.e., error, but one may also interpret the data to suggest that the rater bias is an important factor in the deviations where there is but one rater to judge a student's performance . . . .[24]

Medley and Mitzel in an early study supplemented classroom observation with video tape recordings to obtain a more meaningful record of teacher performance and a measure of student teacher growth.[25] Their instrument (OScAR) and their techniques for reporting teacher behavior have been influential in designing this study, in building the Secondary Student Teacher Performance Profile (SSTPP), and in writing the Guide for Assessment of Secondary Student Teacher Performance (Tentative).

Suggestions for organizing the behavior of teachers in the classroom were gained from Marie Hughes' classification of teaching into seven functions.[26] Her conclusion that teachers' behavior patterns are stable through time was viewed as an hypothesis to be examined. Her model pattern for teacher behavior which may be relevant for elementary schools may have limited implications for this study which is restricted to secondary classrooms.

Flanders' work on interaction analysis suggests the need for careful training of observers.[27]

The one hundred research projects which constitute Ryans' study of teacher characteristics provided a wealth of suggestions for observing teacher performance.[28]

[24] Ibid., pp. 72-76

[25] H. E. Mitzel, A Behavioral Approach to the Assessment of Teacher Effectiveness (New York: Division of Teacher Education, College of the City of New York, 1957). (Mimeographed.)

H. Schueler, M. J. Gold, and H. E. Mitzel, The Use of Television for Improving Teacher Training and for Improving Measures of Student-Teaching Performance, Phase I--Improvement of Student Teaching (New York: Hunter College of the City University of New York, 1962). U.S. Office of Education, Government No. 730035. (Mimeographed.)

[26] Marie Hughes, Development of the Means for the Assessment of the Quality of Teaching in Elementary Schools (Salt Lake City: University of Utah Press, 1959).

[27] N. A. Flanders, "Interaction Analysis in the Classroom: A Manual for Observers" (unpublished manuscript, University of Michigan, Ann Arbor, 1960).

[28] David Ryans, Characteristics of Teachers: Their Description, Comparison, and Appraisal (Washington, D.C.: American Council on Education, 1960).

Under a grant by the Fund for the Advancement of Education, Berry studied the question, "Does completion of education courses prescribed for certification make any difference in the effectiveness of beginning teachers?" Each of 76 first-year teachers having none or only some of the prescribed sequence of education courses was paired with the best "match" that could be found from a pool of 343 certified beginning teachers (age, quality-point average, college major, teaching assignment, extent of work in education). A modified Ryans' Classroom Observation Record was used to observe each teacher five times during the year: twice by professional educators, twice by other professionals (civil engineer, lawyer, physician, social worker) and once by a general observer (a former school superintendent) who visited all teachers participating in the study. The observers were not informed of, and in general were not aware of, the specific problems being investigated nor did they know about the training backgrounds of the teachers observed. The fully certified beginning teachers who had completed the prescribed courses in education were, as a group, consistently and significantly rated to be more effective than the provisionally certified teachers who lacked all or part of the sequence in education courses. The observers from professional fields other than education recognized the superiority of the fully prepared teachers to about the same degree that the educators did.[29]

Bellack and Davitz used a tape recorder and verbatim transcriptions of four class lessons in 15 different classes in social studies to study the teaching process through analysis of the linguistic behavior of teachers and students in the classroom. The study focused on language as the main instrument of communication in teaching. The data were coded according to certain pedagogical moves (structuring, soliciting, etc.) and the results were summarized according to percentages of the various activities.[30]

Many suggestions were obtained from the summary of contemporary research on teacher effectiveness by Biddle and Ellena.[31]

---

[29] John R. Berry, Professional Preparation and Effectiveness of Beginning Teachers (Coral Gables, Florida: Graphic Arts Press, 1960).

[30] Arno Bellack and Joel R. Davitz, The Language of the Classroom (New York: Institute of Psychological Research, Teachers College, Columbia University, 1963).

[31] Bruce J. Biddle and William S. Ellena, Contemporary Research on Teacher Effectiveness (New York: Holt, Rinehart and Winston, 1964).

Several studies which were not used in designing the research program proved helpful as the project progressed. The work of Bush, Allen and Boyan at Stanford on the Experimental Secondary Teacher Education Program suggested useful procedures.[32] Hilda Taba's experiment with teaching for thinking helped shape the staff's perceptions.[33] The later reports of Smith's study of the logic of teaching proved helpful.[34] An interesting compilation of various systems for observing classrooms has been prepared by Simon and Boyer.[35] Studies by Yee,[36] Cyphert,[37] Fall,[38] Mars,[39] and Turner[40] proved helpful in many ways. Medley's subsequent work as presented in recent publications and discussed in personal conferences served to sharpen the staff's perception of the problems of defining teacher behavior.[41]

---

[32] Bush, Boyan, and Allen, op. cit.

[33] Taba, Levine, Elzey, op. cit.

[34] B. Othanel Smith, et al., A Study of the Strategies of Teaching, USOE Cooperative Research Project No. 1640 (Urbana, Ill.: Bureau of Educational Research, University of Illinois, 1967).

[35] Anita Simon and E. Gil. Boyer (eds.), Mirrors for Behavior--An Anthology of Classroom Observation Instruments, under the provisions of Title IV ESEA of 1965 in cooperation with the USOE Research Contract OEC 1-7-062867-3053 (Philadelphia: Research for Better Schools, Inc., 1967).

[36] Albert H. Yee, The Student Teacher Trend, research report to the U.S. Office of Education (Austin, Tex.: University of Texas, 1967).

[37] Fredrick R. Cyphert and Ernest Spaights, An Analysis and Projection of Research in Teacher Education, USOE Cooperative Research Project No. F-015 (Columbus, Ohio: Ohio University Research Foundation, 1964).

[38] Charles R. Fall, et al., Individualized Instruction in Teacher Education. Parts I and II and The Maryvale Project (Buffalo, N.Y.: Inter-University Project One at State University of New York at Buffalo, Cornell University, The University of Rochester, and Syracuse University with the support of the Ford Foundation, 1948).

[39] Walter Jerry Mars, "A Study Comparing Two Ways of Training Students to Observe the Classroom Teacher's Verbal Behavior" (Doctoral dissertation, Syracuse University, Syracuse, N.Y., 1964).

[40] Richard L. Turner, Problem Solving Proficiency Among Elementary Teachers, USOE Cooperative Research Project No. 1262 (Bloomington, Ind.: Institute of Educational Research, 1964).

Research Related to Criteria To Be Used in Evaluating Student Teacher Behavior

Criterion I --Understanding, Friendly vs.
Egocentric, Aloof
Criterion II --Planned, Organized, Responsible vs.
Unplanned, Disorganized, Irresponsible
Criterion III--Stimulating, Imaginative, Surgent vs.
Dull, Routine, Unimaginative

The most extensive study of teachers to date, the Characteristics of Teachers study directed by Ryans, had many implications for this research.[42] The purpose of Ryans' research was to find the traits and behaviors of classroom teachers that are significantly related to teacher effectiveness.

During the more than six years of the study, approximately 100 separate research projects were carried out, and more than 6,000 teachers in 1,700 schools and about 450 school systems participated in the research. Many of the studies involved classroom observation by carefully trained and retrained observers to discover patterns of teacher behavior.

In the teacher characteristics study, the observer, after observing a class for about 50 minutes, made assessments on each of 25 dimensions of the teacher's behavior and then recorded his assessments in the form of ratings on seven-point scales.

As a result of this direct observation and assessment of teacher classroom behavior and subsequent statistical analyses of the measurement data, several interdependent patterns of teacher behavior were suggested. Three in particular appeared to stand out in separate factor analyses of both elementary and secondary teacher data:

Pattern X--warm, understanding, friendly vs.
aloof, egocentric, restricted teacher behavior

---

[41] Donald M. Medley, "The Language of Teacher Behavior: Communicating the Results of Structured Observations to Teachers" (paper presented at Annual Meeting of American Educational Research Association, February, 1967). (Mimeographed.)

Donald M. Medley and Donald L. Lantz, "Classroom Application of Teacher Behavior Research" (paper presented at the Annual Meeting of the American Educational Research Association, February, 1967). (Mimeographed.)

Donald M. Medley, Joseph T. Impellittere, and Lou H. Smith, "Coding Teachers' Verbal Behavior in the Classroom," a manual for users of OScAR 4V. (Mimeographed copy provided by Medley.)

[42] Ryans, op. cit.

Pattern Y--responsible, businesslike, systematic vs.
evading, unplanned, slipshod teacher behavior

Pattern Z--stimulating, imaginative, surgent vs.
dull, routine teacher behavior.

Ryans does not assert that all teacher behavior falls into one of these three patterns suggested by the factor analyses. He does, however, insist that practical experience as well as empirical data indicates that these are three of the principal areas involved in interpersonal relations, and that they might well be given basic consideration in the theory of teacher behavior. These three patterns of teacher behavior which stood out in separate factor analyses of observational data in the *Characteristics of Teachers* study have been incorporated as categories in the SSTPP and serve as evaluative criteria.

Criterion IV--Perceives Self as Competent vs. Perceives Self as Less Than Adequate

In a publication of the Association for Supervision and Curriculum Development, four outstanding college professors, who have devoted much time to research in the area of the adequate personality, agreed on some striking statements concerning the importance of a positive view of self.[43] According to these authors, people who are relatively adequate and self-actualizing have positive views of self. They tend to see themselves as liked, wanted, worthy, and able to do a good job of whatever they undertake. On the other hand, people who are relatively inadequate tend to have a low estimate of their ability to deal with problems. They tend to view themselves as unliked, unwanted, and unlikely to succeed.

In concluding their chapter "The Positive View of Self," they stated that people who are not comfortable with children or who cannot create an atmosphere which encourages the development of more positive selves should be encouraged to seek occupations other than teaching. Furthermore, they state that teachers who destroy the selves of children should be helped to develop more positive views of self or be removed from the profession. Teachers and children, they explain, can be helped to develop more adequate selves in an atmosphere in which they experience success which fosters a positive view of self.[44]

In another chapter of the same yearbook, Combs stated that there are four characteristics of the perceptual field of the truly adequate person. The characteristics that he listed are: "(1) a positive view of self, (2) identification with others, (3) openness to experience and acceptance, and (4) a rich and available perceptual field."[45] The conclusion to be drawn from this and other sections of the book is that a positive view of self is absolutely essential to adequate personality. However, a positive view of self alone does not always assure an adequate person.

---

[43] Arthur W. Combs (ed.), *Perceiving, Behaving, Becoming*, Yearbook, 1962 (Washington, D.C.: Association for Supervision and Curriculum Development, 1962), p. 100

[44] *Ibid.*, p. 117

In a survey of the research concerning self-perceptions, Bills found that most college students could be divided into three approximately equal groups based on their perceptions of themselves and their peers. The three groups were:

> (1) People who accept themselves and who believe that other people in their peer group are equally or more accepting of themselves . . . .
>
> (2) People who are rejecting of themselves but who believe that other people in their peer group are more accepting of themselves . . . .
>
> (3) People who accept themselves but who believe that other people in their peer group are not as accepting of themselves.[46]

He found that 93.1 percent of the college students and 90.9 percent of the high school seniors studied could be included in the three groups listed above. He also found research to indicate that when superintendents are asked to judge the success of principals, supervisors, supervising teachers and assistant superintendents, they designated those of the first type listed above as most successful, those of the second type as next most successful, and those ot the third type as least successful.

The research by Bills would seem to indicate that the persons who hold low estimates of themselves and high estimates of others constitute the middle group as far as rated success in leadership positions is concerned. People who hold high estimates of themselves and high estimates of others are generally rated higher than those who hold low estimates of themselves. In view of Bills' conclusions and those reached by other authors, the relationships between this criterion of student teacher performance and other criteria included in this investigation will be analyzed. It may very well be that this criterion is positively related to good teaching only when combined in certain ways with other criteria.

Combs suggests a model for a teacher education program which is built around the self-concept of the teacher.[47]

---

[45] Ibid., p. 51

[46] Robert E. Bills, About People and Teaching, Bulletin of the Bureau of School Service, Vol. XXVIII, No. 2 (Lexington: University of Kentucky, December, 1955), p. 20.

[47] Arthur W. Combs, The Professional Education of Teachers (Boston: Allyn & Bacon, 1965).

Criterion V--Has Mastery of Facts and Organizing Principles of Field vs. Has Only Minimum Knowledge of Field

This criterion suggests that the student teachers who rate high have more than a superficial acquaintance with the teaching subject. It suggests that the student understands the important concepts of the subject and the rationale upon which it is developed. A student teacher who ranks high in this respect would recognize the most important and significant knowledge in his field--concepts, generalizations, relationships, and speculation on future developments--instead of isolated factual knowledge.

This concept was the focus of the Woods Hole Conference as reported by Bruner in his book The Process of Education. The organizing principles, called "structure" by Bruner, are the basic ideas of how concepts in a subject field are related.[48]

Hunt and Metcalf stress this when they discuss the necessity for generalizations in teaching. A true generalization is timeless in quality, referring to past, present, and future. Discrete or single insights or items or information do not by themselves have meaning or usefulness. A fact can function in thought only when it comes to have evidential character . . . encompassing specifics in the general. The only usable--and therefore worthwhile--"background" consists of if-then-always generalizations. The only role which facts can play in meaningful learning is to function as evidence in a conceptualization. It should therefore be clear that teaching . . . must lead to generalization.[49]

More recently Broudy, Smith and Burnett have pointed out that different levels of factual knowledge exist. They indicate that some factual details are enormously important for both interpretive and applicative uses while others are of minor import. Another dimension is injected as they indicate that the context in which facts are used makes a great deal of difference as to what facts are relevant. What is important to a particular person or in a particular situation may not be as important to other people or in different situations.[50]

---

[48] Jerome Bruner, The Process of Education (Cambridge, Mass.: Harvard University Press, 1962).

[49] Maurice P. Hunt and Lawrence E. Metcalf, Teaching High School Social Studies (New York: Harper and Brothers, 1955), pp. 37-41.

[50] Harry S. Broudy, B. Othanel Smith, and Joe R. Burnett, Democracy and Excellence in American Secondary Education (Chicago: Rand McNally, 1964), pp. 19-21.

Criterion VI --Communicates Well and Empathetically vs.
Communicates Ineffectively and Perfunctorily
Criterion VII--Classroom Discourse Characterized by Reasoning and Creative Thinking vs.
Classroom Discourse Characterized by Simple Recall

The research studies mentioned by Smith, Hughes, Flanders, Ryans, and Bellack are especially relevant for studying the teacher's linguistic behavior and classroom discourse. Aschner's work in analyzing classroom discourse is helpful.[51]

Criterion VIII--Directs Attention to the Logical Operations in Thinking vs.
Fails to Call Attention to the Logical Operations in Thinking

This criterion grows out of the often-expressed need for a methodology for reaching consensus in our democratic society. One necessary skill is the ability and willingness to recognize the value structure which underlies the positions taken by different individuals. Also involved is the ability to apply the appropriate logical analysis to the different positions. The qualities called for in Criterion VIII were explored well in the Twenty-eighth Yearbook of the National Society of College Teachers of Education.[52] Wilson gives a relatively simple analysis of the semantic and logical aspects of this category.[53]

Criterion IX--High Professional Potential vs.
Low Professional Potential

The recent study on the content of a teacher education program by LaGrone has many suggestions for Criterion IX.[54]

---

[51] Mary J. Aschner, The Analysis of Classroom Discourse--A Method and Its Uses (unpublished doctoral dissertation, University of Illinois, Urbana, 1958).

[52] Kenneth D. Benne, George E. Axtelle, B. Othanel Smith, and R. Bruce Raup, The Discipline of Practical Judgment in a Democratic Society, 1943. (Reprinted as The Improvement of Practical Intelligence. New York: Harper, 1950.)

[53] Wilson, op. cit.

[54] Herbert LaGrone, A Proposal for the Revision of the Pre-Service Professional Component of a Program of Teacher Education (Washington, D.C.: American Association of Colleges for Teacher Education, 1964).

CHAPTER III*

DATA COLLECTION AND ANALYSES

The reliability of each of the nine criteria of the SSTPP was analyzed separately and also studied under five experimental conditions. The five conditions, called treatments, referred to the different ways of data collection; the college supervisor was the experimental or treatment unit.

Data Collection

In treatment 1, each college supervisor observed a student teacher on two visits, with the restriction that no student teacher was observed by more than one college supervisor. The college supervisors on each of two visits rated the student teacher's performance in terms of the nine criteria of the SSTPP. At least two weeks elapsed between visits to the same student teacher.

Treatment 2 required that a college supervisor observe a student teacher on four visits, with the same restriction that no student teacher was observed by more than one college supervisor. On each of the four visits, the college supervisor rated the performance of the student teacher on the nine criteria. The time between visits was at least two weeks.

In treatment 3, the college supervisors were divided into teams, with two college supervisors on a team. The team of two college supervisors then observed a student teacher on one visit. During the visit, each member of the team rated the student teacher independently. Student teachers were rated only once by a team and no student teacher was observed by more than one team.

Treatment 4 was a variation of treatment 1. A college supervisor observed and rated the student teacher on two different occasions during the same day. That is, the college supervisor rated the student teacher's performance during a morning class and later that same day rated his performance in an afternoon class. Moreover, no student teacher was observed by more than one college supervisor.

Treatment 5 was a modification of treatment 3. Treatment 5 was added to test the effects of more specific definition of the points on the criterion scale and the effects of a forced choice for criteria VII and VIII. A team of two college supervisors observed and rated the student teacher on one visit.

---

*This chapter was written by Howard Rogers, a member of the staff of the Institute for Research in Human Behavior, Indiana State University.

For treatments 2, 3, 4, and 5, student teachers were randomly selected from the population. Treatment 1 included all student teachers supervised by the Division of Teaching.

Purpose of the Treatments

The treatments were so designed that by comparing them, certain specific questions could be answered. The treatments differed in the type of reliabilities gathered, the number of visits, and the time lapsed between visits. Table 3.1 summarizes these differences. Treatments 1, 2, and 4 generated intra-observer reliabilities and tested the question: How much did the ratings of the same observer show agreement?

Treatment 3 generated inter-observer reliabilities and tested the question: How much did the ratings of two observers, judging the same student teacher, agree?

TABLE 3.1

TREATMENT CLASSIFICATION

Treatment	Type of Reliability	Number of Visits Per Student Teacher	Time Between Visits
1	Intra-Observer	2	2 weeks
2	Intra-Observer	4	2 weeks
3	Inter-Observer	1	- - - -
4	Intra-Observer	2	same day

Comparing treatments 1 and 2 showed the effects of number of visits on their respective reliabilities, since the treatment designs differed only in respect to the number of visits. Comparing treatments 1 and 4 revealed the effects of the time between visits on the reliabilities.

Technique of Analysis

The reliability of the nine criteria of the five treatments was determined by means of analysis of variance (ANOVA). For multifacet data, Medley and Mitzel[1] suggested that the ANOVA was the appropriate way of determining reliability. In estimating reliability from the ANOVA, reliability was defined as the ratio of the true population variance to the total population variance. This definition means that any contribution to the variance by a facet, except the one in which the experimenter was interested, was considered extraneous (error variance) and, in accordance with the definition, contribution by extraneous facets reduced reliability. The ANOVA not only allowed the experimenter to ascertain the amount of variance contributed by each facet, but it also permitted the experimenter to analyze variations in reliability with changes in the sample size of any facet. Reliability determined from the ANOVA could be drastically different from reliability determined from a product-moment correlation. Medley and Mitzel stated:

---

[1] Donald M. Medley and H. E. Mitzel, "Measuring Classroom Behavior by Systematic Observation," N. L. Gage (ed.), Handbook of Research on Teaching, Chapter 6. American Educational Research Association (Chicago: Rand McNally, 1963).

> Reliability [from the ANOVA definition] can be low even though observer agreement is high for a number of reasons. For example, observers might be able to agree perfectly on the number of seats in a room, yet if the number of seats in all rooms is equal, or nearly so, the reliability of seat counts as a measure of differences between classes will be zero. Near-perfect agreement could also be reached about the number of boys in a room wearing red neckties; but if every boy changed the color of his tie every day, the reliability of these counts would be zero.[2]

For reliability determined from the ANOVA to be high, the variance due to "the facet of interest" must be large relative to other sources of variance. However, an advantage of the ANOVA is its use in suggesting what facets should be deleted and also added to a future study. For a detailed explanation of estimating reliability by the analysis of variance, the reader is referred to the cited study by Medley and Mitzel and to Gleser *et al*.[3]

Reliabilities of the Treatments

Factors in the treatments could be subjected to various interpretations; however, it must be recognized that each college supervisor's ratings were considered independent. Also, since treatments 1, 2, and 4 used the same statistical design, the three treatments are discussed first and some of their data are presented in the same tables for purposes of comparison.

Intra-observer Reliability. The reliabilities of criteria VII and VIII are not shown for treatments 1, 2, 3, and 4. The reason for their deletion was that at least 30 per cent of the ratings for these two criteria were "suspended judgments"; the college supervisor marked the columns N (non-sufficient evidence for judgment) on these occasions.

---

[2] *Ibid*., p. 310.

[3] G. C. Gleser, L. J. Cronbach, and N. Rajaratnam, "Generalizability of Scores Influenced by Multiple Sources of Variance," *Psychometriker*, Vol. 30, 1965.

Attempts to incorporate the "suspended judgments" into the statistical analysis were unsatisfactory. Moreover, the number of suspended judgments was further evidence of the difficulty of rating student teachers in the two criteria and this fact could be considered proof of their "unreliability."

In treatments 1, 2, and 4, the experiment was a partially hierarchical design. Student teachers were nested under college supervisors, and visits were crossed with college supervisors and student teachers. However, it should be noted that even though the three treatments employed the same statistical design, the levels of a factor (visits) and/or the definition of a factor (visits) were different in the three treatments. The statistical model for the three separate treatments was:

$$Xcsv = U + Ac + Bs(c) + Cv + ACcv + ABCs(c)v$$

where c, s, and v were college supervisors, student teachers, and visits, respectively.

Table 3.2 and Table 3.3 summarize the pertinent information needed to estimate reliability. First, Table 3.2 presented the sources of variation, the degrees of freedom associated with each source, the expected mean square and the obtained mean square. Estimations of the variance components were derived from information contained in Table 3.2. The derived estimates were obtained by setting the expected mean square components equal to their respective obtained mean squares and solving the resulting set of linear equations in terms of each variance component; that is, $\sigma^2 sv(c) = S^2 sv(c)$ and $\sigma^2 sv(c) + s\sigma^2 cv = S^2 cv$, etc. By solving each equation in turn, all the population variances (σ^2) can be written in terms of the obtained mean square values. The formula for estimating variance components from the obtained mean squares is shown in Table 3.3.

To obtain the reliability of the criteria, the total population variance and true population variance must be estimated. The estimation of the total population variance depended upon the experimental design. For the three treatments, the total population variance equaled:

$$\sigma^2 c + \frac{1}{v}\sigma^2 cv + \frac{1}{s}\sigma^2 s(c) + \frac{1}{sv}\sigma^2 s(c)v$$

and the true population variance equaled:

$$\sigma^2 c$$

(The design and analysis of the data were equivalent to design 4 generalized to Universe A in Gleser *et al.*[4])

In this design, the variance due to visits dropped out, because the effects of the visit factor were the same for all college supervisors. The reliability formula for the three treatments was:

$$\sigma^2 c/\sigma^2 c + \frac{1}{v}\sigma^2 cv + \frac{1}{s}\ \sigma^2 s(c) + \frac{1}{sv}\ \sigma^2 s(c)v$$

---

[4] *Ibid.*

TABLE 3.2

ANOVA DESIGN FOR TREATMENTS 1, 2, AND 4

Source of Variation	Treatment 1 DF	Treatment 2 DF	Treatment 4 DF	Expected Mean Square	Obtained Mean Square
A College Supervisors	(c-1) (11-1)	(4-1)	(4-1)	$\sigma^2 s(c)v + s\sigma^2 cv + v\sigma^2 s(c) + sv\sigma^2 c$	$S^2 c$
BWA Student Teachers	c(S-1) 11(28-1)	4(5-1)	4(5-1)	$\sigma^2 s(c)v + v\sigma^2 s(c)$	$S^2 s(c)$
C Visit	(V-1) (2-1)	(4-1)	(2-1)	$\sigma^2 s(c)v + s\sigma^2 cv + cs\sigma^2 v$	$S^2 v$
AC Coll. Sup. X Visit	(c-1)(v-1) (11-1)(2-1)	(4-1)(4-1)	(4-1)(2-1)	$\sigma^2 s(c)v + s\sigma^2 cv$	$S^2 cv$
(BWA)C Residual	c(S-1)(V-1) 11(28-1)(2-1)	4(5-1)(4-1)	4(5-1)(2-1)	$\sigma^2 s(c)v$	$S^2 s(c)v$

TABLE 3.3

FORMULA FOR DERIVATION OF VARIANCE COMPONENTS

FOR TREATMENTS 1, 2 AND 4

$$\begin{aligned}
\sigma^2 c &= \frac{1}{vs}(S^2 c - S^2 cv - S^2 s(c) + S^2 s(c)v) \\
\sigma^2 s(c) &= \frac{1}{v}(S^2 s(c) - S^2 s(c)v) \\
\sigma^2 v &= \frac{1}{cs}(S^2 v - S^2 cv) \\
\sigma^2 cv &= \frac{1}{s}(S^2 cv - S^2 s(c)v) \\
\sigma^2 s(c)v &= S^2 s(c)v
\end{aligned}$$

The reliabilities for treatments 1, 2, and 4 are shown in Table 3.4. For treatment 1, the reliability of each criterion was based on eleven college supervisors, each observing twenty-eight students on two visits. For treatment 2, the reliability of each criterion was based on four college supervisors, each observing five students on four visits. For treatment 4, the reliability of each criterion was based on four college supervisors, each observing five students on two visits. Basically, the reliabilities of treatments 1 and 2 were moderately low, while that of treatment 4 was extremely low. (Adequate reliability was considered to be .80 or higher.) Over the treatments, there was no constant trend of one criterion's reliability being higher than that of another.

TABLE 3.4

RELIABILITY OF SEVEN CRITERIA

TREATMENTS 1, 2 AND 4

Criterion		Reliability	
	Treatment 1	Treatment 2	Treatment 4
I	.69	.66	.23
II	.65	.81	.00
III	.32	.69	.00
IV	.68	.73	.00
V	.66	.21	.00
VI	.50	.82	.00
IX	.39	.66	.00

Reasons for Reliability Results. The magnitude of variance components afforded a method of comparing the three treatments, since the estimated variance components were independent of sample size of the facets. An examination of the variance components in Table 3.5 revealed certain consistencies and differences between the three treatments.

In treatments 1 and 4, the magnitude of the variance components showed consistent trends over the seven criteria. $\sigma^2 s(c)v$ (student-teachers-within-a-college-supervisor component) had the largest absolute magnitude for the seven criteria and also its relative magnitude compared to $\sigma^2 v$ (visits component), $\sigma^2 cv$ (visits x college supervisors component), and $\sigma^2 c$ (college supervisors component) was large. The relative and absolute magnitudes of the last three estimated variance components were quite small. The relative magnitude of $\sigma^2 s(c)$ meant the variation of student teachers under (within) a college supervisor was greater than the variation between college supervisors ($\sigma^2 c$). Medley and Mitzel reported similar results in their observational study of classroom behavior: the variations within a classroom were always greater than variations between classrooms.[5] Although the magnitude of $\sigma^2 s(c)v$ was smaller than $\sigma^2 s(c)$ over the seven criteria, in most cases, its relative magnitude compared to $\sigma^2 v$, $\sigma^2 cv$ and $\sigma^2 c$ was large. The size of $\sigma^2 s(c)v$ (residual component) showed that there were factors affecting the ratings which were not isolated in this experiment.

In treatment 2, the criteria seemed to divide into two groups. The relative magnitudes of the variance component in criterion V were similar to those of treatments 1 and 4. Criteria I, II, III, IV, VI, and IX, however, showed another pattern. As in treatments 1 and 4, the relative magnitudes of $\sigma^2 s(v)$ and $\sigma^2 s(c)v$ compared to $\sigma^2 v$ and $\sigma^2 cv$ were large; but for these six criteria of treatment 2, unlike the other treatments, the relative magnitude of $\sigma^2 c$ was large.

The magnitude of $\sigma^2 s(c)$ meant that the college supervisors were giving differential ratings to their student teachers. That is, the college supervisors did not assign the same rating to their student teachers; however, in treatments 1 and 4, the difference between the average ratings of the college supervisor was small compared to the differences between their average ratings of student teachers. In treatment 2, the difference between college supervisors on six criteria was higher than in the other two treatments, indicating that the reliability for the six criteria was higher than those found in the other treatments. The residual error was large across criteria and treatments. This suggested that more factors should have been included in the statistical model.

Examination of the variance components raised several related questions. First, why was $\sigma^2 c$ (college supervisors components) in treatment 4 so small (zero in five out of the six criteria) when compared to $\sigma^2 c$ in treatment 2? The answer could be seen in the derivation of $\sigma^2 c$. The

---

[5] *Op. cit.*

TABLE 3.5

ESTIMATION OF VARIANCE COMPONENTS IN TREATMENT 1, 2 AND 4

	Criterion I			II			III			IV			V			VI			IX		
	Treatment			Treatment			Treatment			Treatment			Treatment			Treatment			Treatment		
Source of Variance	1	2	4	1	2	4	1	2	4	1	2	4	1	2	4	1	2	4	1	2	4
$\sigma^2 c$	.07[a]	.14	.06	.07	.55	.00	.03	.38	.00	.09	.42	.00	.08	.08	.00	.05	.44	.00	.03	.26	.00
$\sigma^2 s(c)$	.55	.19	.95	.52	.50	1.03	.80	.42	.87	.65	.65	.49	.73	1.14	.95	.69	.36	1.11	.75	.45	1.36
$\sigma^2 v$	.01	.07	.00[b]	.00	.04	.04	.01	.00	.00	.04	.03	.00	.01	.00	.10	.02	.02	.00	.01	.00	.03
$\sigma^2 cv$	.02	.02	.00	.02	.02	.00	.03	.25	.00	.02	.01	.00	.04	.16	.00	.04	.02	.15	.03	.11	.00
$\sigma^2 s(c) \cdot v$	.34	.55	.15	.46	.50	.51	.55	.49	.60	.52	.38	.24	.42	.38	.33	.55	.47	.39	.36	.36	.35

[a] Values to four decimal places used in computation.

[b] Negative variance equals zero.

formula was:

$$\sigma^2c = \frac{1}{vs}(S^2c - S^2s(c) - S^2ac + S^2s(c)v)$$

In treatment 4, the magnitude of $S^2s(c)$ (obtained mean square of student-teachers-within-a-college supervisor) was so much greater than the other obtained mean squares, including S^2c, that the sum of the four obtained mean squares was negative (which was transformed to zero, since no negative variance could occur). The magnitude of $\sigma^2s(c)$ in treatment 2 was never greater than σ^2c.

But this difference in the relative size of S^2c and $S^2s(c)$ in the two treatments generated another problem. Since in both treatments, the college supervisor rated student teachers from the same population, why should the variance of the student teachers and college supervisors in the two treatments be so different? Several alternative explanations were examined. The results of the examination revealed no consistent reason why the variance components should be so different in the treatments, except chance variations. Since the intercorrelations between the seven criteria were approximately .5, any random difference between treatments would be expected to produce differences in all the criteria.

A second related question was: Why were the reliabilities in treatment 4 so low? Since σ^2c was zero or relatively small in all cases, the reliability, as a consequence of the definition used, must also be zero or insignificant.

Since the error variation was so large, a product-moment correlation between two visits was performed. The correlation for treatments 1 and 4 was between the first and second visit. For treatment 2, it was between the first and third visit, since the time lapse between them was equivalent to the time lapse between the first and second visit in treatment 1. Table 3.6 shows the results. The analysis was performed on the combined data of all the college supervisors. The result showed that the correlation between visits was highest in treatment 4 and approximately equivalent in treatments 1 and 2, the exception being criterion VI. For Criterion VI, treatment 2 was slightly higher than treatment 4. The correlations suggested that the shorter the interval between visits was, the higher the product-moment correlation between the visits would be. This fact might mean that with a small time interval, the college supervisors remembered their previous ratings and on the second visit, attempted to minimize--whether consciously or unconsciously--any difference between the two ratings. Although pooling the data of all the college supervisors, the correlations could be used as a crude measure of correlation between the two visits.

As had been previously mentioned, an advantage of using an ANOVA for determining reliability was that the estimated variance components could be used to calculate reliabilities based on different sample sizes. The purpose of varying the sample size of the facets was to allow the experimenter to estimate the adequacy of the reliability for various conditions: if the reliability of a criterion was low, the experimenter, using the estimates of various variance components from the original study, could determine which facets should be increased to obtain higher reliability.

TABLE 3.6

PRODUCT-MOMENT CORRELATIONS OF CRITERION RATINGS
IN TREATMENTS 1, 2 AND 4

Criterion	Treatment 1 N = 499	Treatment 2 N = 20	Treatment 4 N = 23
I	.63	.59	.80
II	.62	.61	.69
III	.59	.42	.68
IV	.60	.61	.77
V	.64	.60	.69
VI	.53	.62	.57
VII	.42	.49	.59
VIII	.41	.39	.73
IX	.68	.62	.80

From the reliability formula, $r = \sigma^2 c/\sigma^2 c + \frac{1}{v}\sigma^2 cv + \frac{1}{s}\sigma^2 s(c) + \frac{1}{sv}\sigma^2 s(c)v$, it could be easily shown that increasing the number of student teachers and the number of visits would increase the reliability, unless the estimated $\sigma^2 c$ equaled zero. In that case, increasing sample size could not affect reliability, which would always be zero.

To gauge the adequacy of the reliability in the three intra-observer treatments, a projected study, using the original estimated variance components, was performed. In the projected study, each college supervisor observed the same two student teachers on ten visits. The choice of the sample size was based on the idea that the reliability of a minimal sample of student teachers should be high and the way to accomplish this end was to increase the number of visits (within reason). The results of the projected study are shown in Table 3.7. The reliabilities of the seven criteria were low. The reason for the low reliabilities was that the magnitude of student teacher variance was unaffected by changes in the number of visits and thus was relatively large.

The size of some variance components suggested that they were not significantly different from zero. An F test was performed on the data to ascertain if some of the sources of variance could be pooled. To apply the F test, it was assumed that all the factors and the residual variation were normally distributed. Tables 3.8, 3.9 and 3.10 show the results of the F test for treatments 1, 2 and 4, respectively. For the college-supervisors factor, a quasi F ratio was used. At this stage, the value of the F test could be its use as a guide in suggesting what

TABLE 3.7

PROJECTED RELIABILITY OF THREE TREATMENTS

Criterion	Treatment 1	Treatment 2	Treatment 4
I	.19	.56	.11
II	.20	.66	.00
III	.06	.62	.00
IV	.20	.55	.00
V	.17	.11	.00
VI	.12	.69	.00
IX	.07	.51	.00

TABLE 3.8

ANOVA FOR TREATMENT 1

Source of Variance	DF	Criterion I MS	Criterion I F	Criterion II MS	Criterion II F	Criterion III MS	Criterion III F	Criterion IV MS	Criterion IV F	Criterion V MS	Criterion V F	Criterion VI MS	Criterion VI F	Criterion IX MS	Criterion IX F
A	10	6.01	3.17**	6.03	2.87*	4.48	1.47	7.11	3.10**	6.35	2.97**	6.00	2.00	4.64	1.63
BWA	297	1.45	4.26**	1.50	3.26**	2.15	3.91**	1.82	3.50**	1.88	4.48**	1.93	3.51**	1.86	5.17**
C	1	2.89	3.71	1.11	1.05	3.92	2.72	12.28	12.41**	2.91	4.28	9.03	5.57*	5.34	3.99
AC	10	.78	2.29*	1.06	2.30*	1.44	2.62**	.99	1.90*	.68	1.62	1.62	2.95**	1.34	3.72**
(BWA)C	297	.34		.46		.55		.52		.42		.55		.36	

$*p < .05$
$**p < .01$

TABLE 3.9

ANOVA FOR TREATMENT 2

Source of Variance	DF	Criterion I MS	Criterion I F	Criterion II MS	Criterion II F	Criterion III MS	Criterion III F	Criterion IV MS	Criterion IV F	Criterion V MS	Criterion V F	Criterion VI MS	Criterion VI F	Criterion IX MS	Criterion IX F
A	3	4.16	2.95	13.63	5.24*	10.98	3.23*	11.32	3.76*	7.27	1.26	10.80	5.45**	7.80	2.90
BWA	16	1.30	2.36*	2.48	4.96**	2.16	4.41**	2.96	8.22**	4.95	13.03**	1.90	4.04**	2.14	5.94**
C	3	2.13	3.227	1.33	2.15	.83	<1.00	.97	2.37	1.03	<1.00	.88	1.60	.58	<1.00
AC	9	.66	1.20	.62	1.24	1.73	3.53**	.41	1.14	1.19	3.13**	.55	1.17	.91	2.52**
(BWA)C	48	.55		.50		.49		.36		.38		.47		.36	

*$p < .05$
**$p < .01$

TABLE 3.10

ANOVA FOR TREATMENT 4

Source of Variance	DF	Criterion I		Criterion II		Criterion III		Criterion IV		Criterion V		Criterion VI		Criterion IX	
		MS	F	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F
A	3	2.69	1.28	1.47	<1.00	.27	<1.00	.76	<1.00	2.18	<1.00	2.82	<1.00	1.38	<1.00
BWA	16	2.05	13.66**	2.57	5.03**	2.34	3.90**	1.23	5.13**	2.23	6.76**	2.62	6.72**	3.07	9.03**
C	1	.10	<1.00	1.05	2.06	.03	<1.00	.14	<1.00	2.30	6.76	.11	<1.00	.55	1.62
AC	3	.16	1.07	.28	<1.00	.30	.50	.21	<1.00	.34	1.03	1.14	2.92	.03	<1.00
(BWA)C	16	.15		.51		.60		.24		.33		.39		.34	

**p < .01

factors could be dropped from a model of a future study. Across treatments and criteria, none of the sources of variance were consistently non-significant. Therefore, caution would require that all the sources be retained. Moreover, it would not be advisable to drop factors, until adequate reliability was obtained.

Inter-Observer Reliability

For treatment 3, the appropriate statistical model was:

$$X_{tcs} = U + {}^{A}{}_{t} + {}^{B}{}_{c(t)} + {}^{C}{}_{s(tc)}$$

where t, c, and s are the teams, college supervisors, and student teachers, respectively.

In this design, college supervisors are nested under the team factor and student teachers are nested under both college supervisors and team factors.

Table 3.11 summarizes the pertinent information needed to estimate reliability.

Based on the estimated variance components, the total population variance equaled:

$$\sigma^2(t) + \frac{1}{c}\,\sigma^2 c(t) + \frac{1}{cs}\,\sigma^2 s(ct)$$

and the true population variance equaled:

$$\sigma^2(t).$$

Therefore, the reliability of treatment 3 was:

$$\sigma^2(t)/\sigma^2(t) + \frac{1}{c}\,(t) + \frac{1}{cs}\,\sigma^2 s(ct).$$

The reliability of each variable was based on fifteen teams with two college supervisors on a team and each team observing two student teachers. The results are shown in Table 3.12. (The design and analysis of the data were equivalent to design 7 generalized to Universe A in Gleser *et al.*[6])

Table 3.12 revealed that the inter-observer reliabilities were moderately low. For comparison, the product-moment correlations are shown in Table 3.13. Examination of the variance component estimations in Table 3.14 revealed that $\sigma^2 s(ct)$--the error term--was quite large relative to the other factors. The size of the error term implied that variables other than those considered in the design were influencing the ratings. Since

---

[6] *Op. cit.*

TABLE 3.11

ANOVA DESIGN FOR TREATMENT 3

Source of Variation	DF	EMS	Obtained MS
A Team	(t-1) (15-1)	$\sigma^2 s(tc) + s\sigma^2 c(t) + cs\sigma^2 t$	$S^2 t$
B/A College supervisors within team	t(c-1) 15(2-1)	$\sigma^2 s(tc) + s\sigma^2 c(t)$	$S^2 c(t)$
C/AB Student teachers within team x college supervisor	tc(s-1) 15.2(2-1)	$\sigma^2 s(tc)$	$S^2 s(tc)$
TOTAL	tcs-1 (59)		

$\sigma^2 s(ct)$ was so large, it overshadowed the variance due to $\sigma^2 c(t)$--the college supervisors-within-a-team. Therefore, the size of the variance components suggested that inter-observer reliability was reduced, because differences among student teachers had not been controlled or incorporated into the statistical design. To increase the reliability, either the number of college supervisors or the number of student teachers could be increased. A projected study was performed. Each team was composed of five college supervisors who observed two student teachers. The results are shown in Table 3.15. The reliabilities were higher than before, but most were moderate with only two reaching an adequate level of reliability. It would be possible to increase them, but the cost of adding more college supervisors might be prohibitive.

TABLE 3.12

INTER-OBSERVER RELIABILITIES FOR TREATMENT 3

Criterion	Reliability
I	.60
II	.48
III	.53
IV	.56
V	.34
VI	.45
IX	.46

TABLE 3.13

PRODUCT-MOMENT CORRELATIONS OF CRITERION RATINGS IN TREATMENT 3

Criterion	Reliability
I	.15
II	.51
III	.58
IV	.55
V	.52
VI	.55
IX	.48

TABLE 3.14

ESTIMATION OF VARIANCE COMPONENTS IN TREATMENT 3

Criterion	I	II	III	IV	V	VI	IX
σ^2t	.35	.38	.78	.64	.28	.36	.32
$\sigma^2c(t)$	.00	.00	.00	.00	.00	.00	.00
$\sigma^2s(tc)$	.96	1.63	2.79	1.99	2.11	1.74	1.49

TABLE 3.15

PROJECTED STUDY FOR TREATMENT 3

Criterion	Reliability
I	.80
II	.70
III	.74
IV	.80
V	.57
VI	.68
IX	.68

Examination of the F ratios in Table 3.16 suggested that the college supervisors-within-a-team factor was not significant and might be dropped from future models. However, the deletion could be done after another study verified that this factor was insignificant.

Based on the original treatment, it was decided that a possible reason for the moderate reliabilities was that the observers' frame of reference was dissimilar. Therefore, it was felt that a more rigid scale would produce greater conformity in rating student teachers. In treatment 5, the scaling procedure for Criteria VII and VIII was modified. For these two criteria, the endpoints and midpoints were described in detail and specific examples for rating both extremes were identified. Table 3.17 shows the definitions used for Criteria VII and VIII. There were three differences

TABLE 3.16

ANOVA FOR TREATMENT 3

Source of Variance	DF	Criterion I MS	Criterion I F	Criterion II MS	Criterion II F	Criterion III MS	Criterion III F	Criterion IV MS	Criterion IV F	Criterion V MS	Criterion V F	Criterion VI MS	Criterion VI F	Criterion IX MS	Criterion IX F
A	14	2.10	3.04**	1.97	4.28**	3.73	5.92**	3.34	4.18**	1.91	2.39	1.91	3.98**	2.15	2.42
B/A	15	.69	.72	.46	.28	.63	.23	.80	.40	.80	.38	.48	.28	.89	.60
C/AB	30	.96		1.63		2.79		1.99		2.12		1.74		1.49	

**p $<$ .01

TABLE 3.17

REVISED INSTRUCTIONS FOR CRITERION VII

Score as a "1"

If most of teacher's questions are open-ended or thought-provoking.

If teacher frequently translates principles into terms which are in the pupil's fields of experience.

Score as a "4"

If teacher ignores good opportunities for thought part of the time and seizes them part of the time.

Score as a "7"

If students are never asked "why?" or "how?"

If students are not permitted a moment to think of answer.

If no analogies to students' experience are given.

REVISED INSTRUCTIONS FOR CRITERION VIII

Score as a "1"

If teacher makes students aware of the problem and the process for solving it.

If students actually comprehend the problem.

Score as a "4"

If teacher understands the problem and meaning of terms but fails to communicate it to students.

If teacher assumes student will comprehend the process.

Score as a "7"

If teacher never identifies the problem nor defines terms.

2

between the new procedure and the old procedure. One was the elaboration of the midpoint and extremes. Another was that three practice sessions were held in which observers viewed video tapes of student teachers, compared their ratings and discussed the reasons for their different ratings until agreement was reached. Thirdly, the college supervisors were forced to make a judgment on every visit; "suspended" judgments would not be accepted. The remaining criteria in treatment 5 were not modified.

The experimental treatment was similar to treatment 3. A team of two observers rated five student teachers on one visit. Since there was only one team, the statistical design for treatment 3 was inappropriate. The statistical design used in this study was: $X_{sc} = A_s + B_{sc}$, where s represented a student teacher and c represented a college supervisor.

Table 3.18 shows the ANOVA design for treatment 5 and Table 3.19 shows the estimation of variance components.

TABLE 3.18

ANOVA DESIGN FOR TREATMENT 5

Source of Variance	DF	E(MS)	Obtained MS
Students	4	$\sigma^2\text{error} + c\sigma^2 s$	$S^2 s$
Within Students	5	$\sigma^2\text{error}$	

TABLE 3.19

ESTIMATION OF VARIANCE COMPONENTS

$$\sigma^2 s = \frac{1}{c}(S^2 s - S^2\text{error})$$

$$\sigma^2\text{error} = S^2\text{error}$$

In this design, differences in ratings of a single student teacher, σ^2error, were considered a source of error. Differences between student teachers, σ^2s, were considered an estimation of the true characteristic being measured and, therefore, this variance component did not contribute to the error. Thus, the reliability was $\sigma^2 s/\sigma^2 s + \frac{1}{K}\sigma^2\text{error}$. For further detail, Winer has a discussion of this design.[7]

---

[7] B. J. Winer, *Statistical Principles In Experimental Design* (New York: McGraw-Hill, 1962).

Table 3.20 shows the reliabilities. The data supported the idea that a rigid frame of reference produced higher reliability. With the prescription of scale points for Criteria VII and VIII, the reliabilities of these two criteria now became the highest. Therefore, it would seem reasonable that to achieve uniformity of observers' ratings, the rating scale must be anchored, at the endpoints and midpoint, so the observers have a common frame of reference. Also, Criteria I and II now had the lowest reliabilities. The reason for this dramatic decrease was not clear when compared to treatment 3; however, it was possible that with all the emphasis in training on Criteria VII and VIII, the college supervisors may have subsequently ignored Criteria I and II.

TABLE 3.20

RELIABILITIES OF CRITERIA FOR TREATMENT 5

Criterion	Reliability
I	.44
II	.33
III	.91
IV	.87
V	.69
VI	.87
VII	.98
VIII	.94
IX	.89

The detailed analysis of the previous treatment was not performed, because of the limited sample size and because interest was focused only on Criteria VII and VIII. Table 3.21 shows the variance components only for Criteria VII and VIII. The size of the error variance component was insignificant compared to the student teacher variance component, $\sigma^2 s$. However, it must be noted that changes in designs can shift reliabilities drastically and a true replication of treatment 3 might not have produced the same results.

TABLE 3.21

VARIANCE COMPONENT FOR TREATMENT 5

	Criterion VII	Criterion VIII
$\sigma^2 s$	4.08	3.11
$\sigma^2 error$	.20	.39

Discussion of Reliability

Intra-observer Reliability. The results of treatments 1, 2 and 4 revealed that the reliability of college supervisors' ratings of two student teachers was low or moderately low. Even the product-moment correlation (pooled over college supervisors) revealed that over all the treatments none of the criteria consistently obtained a correlation of .80 or greater. This fact meant that either the college supervisors were inconsistent or the student teacher's behavior was variable or both. Obviously, there was no simple way of testing these alternative explanations, since the data showed no systematic trends which could have been attributed to the student teachers or to the college supervisors. However, a repeated measures design (Table 3.22) showed that the college supervisors did not change significantly over time. This evidence probably indicated that the student teachers were variable. That is, the performance of the student teacher probably depended on the nature of the material he was teaching and how the class reacted to him. For example, in treatment 4, where the two visits were on the same day, a student teacher in most cases taught two different classes--subject matter and students--which probably accounted for the low reliability. However, it was still possible that college supervisors' behavior was unsystematically variable.

A comparison of the treatment variations showed significant differences; however, none of the treatments reached an acceptable level of reliability (.80 or greater).

Inter-observer Reliability. The inter-observer reliabilities found in treatment 3 were moderate. The projective study showed that with five college supervisors, the reliabilities of the criteria tended to approach the acceptable range (.80 or greater). However, having five college supervisors observe a student might be impractical in most cases.

When special training and methods were introduced, as in treatment 5, the reliabilities seemed to increase. The importance of treatment 5 was the initial training in which college supervisors were required to discuss their differences until agreement was reached at what the specific rating should be. Before this change, the initial training required a discussion

TABLE 3.22

ANOVA FOR CHANGES OVER TIME

Source of Variation	DF	Criterion I MS	Criterion I F	Criterion II MS	Criterion II F	Criterion III MS	Criterion III F	Criterion IV MS	Criterion IV F	Criterion V MS	Criterion V F	Criterion VI MS	Criterion VI F
Between Subjects	3												
Within Subjects	28												
Conditions (time)	7	1.839	2.046	1.500	1.448	1.982	1.563	.839	.746	1.352	1.543	1.125	.832
Residual	21	.898		1.035		1.267		1.125		.876		1.351	

Source of Variation	DF	Criterion VII MS	Criterion VII F	Criterion VIII MS	Criterion VIII F	Criterion IX MS	Criterion IX F
Between Subjects	3						
Within Subjects	28						
Conditions (time)	7	3.142	.509	7.571	1.154	1.482	1.791
Residual	21	6.166		6.659		.827	

of differences, although the agreement was very general: the student teacher seen on the video tape was rated high, medium or low on a criterion, but nc exact rating, such as "4," would be agreed on. The standardized procedure used with criteria VII and VIII, which previously had been so difficult to rate, now had the highest reliability.

The reliability problem could be interpreted as a special case of discrimination learning. In the same way that a child learned to tell time, college supervisors could be taught to make similar judgments. The child repeated his estimate of the time, until the judgment agreed with his teacher. In treatment 3, no elaborate effort was made to test if the initial frame of references was the same. Treatment 5 elaborated a procedure for initial agreement on criteria VII and VIII and the results were obvious.

Additional Analysis

The reliability analysis of the five treatments raised an additional question: Since the student teacher variance was so large, was the cause of this large variance due to the same systematic changes in the behavior of the college supervisors' responses with the passage of time. Four college supervisors were studied over a four-year period; two ratings per year were randomly selected for four years for each of the college supervisors. The results of the analysis of variance, using the repeated measures design, showed that the time variable had no significant effect on the ratings for any of the nine criteria (see Table 3.22). The results meant that the college supervisors did not change their ratings over time. However, it did not preclude the possibility that the college supervisors were unsystematically changing their ratings, or that the student teachers' behavior was variable. Basically, the reason for the student teacher variance was confounded in this study.

In addition, from an earlier factor analysis, three factors were extracted from the SSTPP. The justification for the extraction of three factors was based purely on rational grounds, since the last factor was insignificant. The factors were a "personality" factor (criteria I, III, IV, VI, and IX), a "logical thinking" factor (criteria VII and VIII), and an "organization-subject-matter" factor (criteria II and V). Since they appeared to be reasonable, factor scores were computed for the three factors and used in the validity studies. A factor score for an individual was the subject's standard score on the sixth criterion times the regression coefficient between the factor and the criterion I, summed over the eight criteria.

Also, the treatment data suggested that the nine criteria were not independent. A factor analysis offered a means of describing these interrelationships more succintly. A factor analysis was performed on the SSTPP (only criteria one through eight), using 1128 cases. The communality estimation option employed in this study was the square of the multiple correlation coefficients. The results in Table 3.23 show that the communalities (common variance) accounted for 58 percent of the total variance (or 4.67/8.00). The first factor accounted for 75 percent of the common variance and the second factor accounted for 22 percent of the common variance. The other two factors were insignificant. The initial factorization revealed that two factors accounted for almost all of the common variance.* Table 3.24 shows the contribution of each criterion (based on the first two factors) to the common variance. The result shows that each test contributes approximately the same amount.

TABLE 3.23

FACTORS OF SSTPP

Characteristic Roots		Percent of Trace
Factor I	3.4904	75%
Factor II	1.0604	22%
Factor III	0.1064	2%
Factor IV	4.6734	1%
	4.6734	100%

TABLE 3.24

COMMUNALITIES OF CRITERIA

Criterion	Communality
I	.4578
II	.5277
III	.6824
IV	.5423
V	.5381
VI	.6040
VII	.5989
VIII	.5996

---

*Using a different group of student teachers three factors were identified. They are described in Chapter IV.

Table 3.25 shows the orthogonal factor matrix. The factor matrix showed that on factor I, the first six criteria had high loadings, while criteria VII and VIII had insignificant loadings ($<.35$). But on factor II, the reverse was true: criteria VII and VIII had significant loadings, while the first six criteria's loadings were insignificant. Moreover, the first six criteria were performance variables: personality (I, III, IV, VI) or organization-of-subject matter variables (II, V). As a consequence of the criteria's definitions, it was reasonable to conclude that factor I was a "general appearance" variable and factor II was a "logical thinking" variable.

TABLE 3.25

ORTHOGONAL FACTOR MATRIX (VARIMAX) OF SSTPP

Variable	Factor 1	Factor 2
Criterion I	0.6709	0.0872
Criterion II	0.7179	0.1106
Criterion III	0.8200	0.0995
Criterion IV	0.7341	0.0574
Criterion V	0.7142	0.1673
Criterion VI	0.7652	0.1355
Criterion VII	0.1274	0.7633
Criterion VIII	0.0954	0.7684

The correlation among the oblique reference vectors was non-significant (.29) and rotation to an oblique reference structure did not appear to affect the factor loadings. Basically, the factors were orthogonal.

In summary, the college supervisor viewed the student teacher in terms of two independent factors. The student teacher's behavior was rated independently along a "general appearance" and along a "logical thinking" dimension.

Moreover, according to the definition, criterion IX was a global variable. It was an overall estimation of the student teacher's potential. Therefore, a stepwise regression analysis of the relationships between criterion IX of the SSTPP as the dependent variable and the other eight categories as the independent variables was performed. There were 499 student teachers in the analysis. The resulting model was:

$$X_9 = -.2232 + .1540X_1 + .2225X_2 + .1729X_3 + .0718X_4 + .2071X_5 + .2027X_6 + .0197X_7 + .0205X_8$$

where X_i was a score on a criterion. The multiple R was .8869, which meant 78.66 per cent of variance was accounted for by the regression equation. Examination of the weights for the criteria revealed that variables 1, 2, 3, 5, and 6 have the highest weights. Except for criterion 5, the rest of the criteria could be characterized as personality variables. The best index of a student teacher's potential was his personality according to college supervisor's judgment. Variables 7 and 8, the "logical thinking" criteria, had the lowest weight in the regression equation. The finding meant that "logical thinking" ability was not a very important element in the college supervisor's rating of the student teacher's potential.

CHAPTER IV

THE TEACHER CLASSROOM ACTIVITY PROFILE

Background

The Teacher Classroom Activity Profile (TCAP), which was designed for one of the "treatments" of the experimental observations, evolved into a major investigation on its own. This is an instrument for recording the way the teacher spends his time in the classroom. The seven categories which finally emerged are: Management--Non-Learning Activities, Management--Learning Activities, Presentation, Conducting Recitation/Drill, Discussion/Random, Conducting Logical Thinking, and Calling Attention to the Process of Thinking. They reflect the bias of this investigation toward the intellectual aspects of behavior.

The seven categories and their definitions for classifying what the teacher does in the classroom were agreed upon after several trial runs and extensive staff discussion.* "Management--Non-Learning" refers to the time spent in housekeeping activities and disciplining. "Management--Learning" refers to time in which the teacher is monitoring the class while the students are engaged in study, work, viewing a film, giving reports, or taking a test--time in which the teacher's role is a passive one. "Presentation" refers to the time the teacher is lecturing, reading, or presenting any kind of information. "Recitation/Drill" is used in the traditional sense of the teacher's asking for brief answers--time in which the teacher is orally discovering whether or not students have prepared their lesson. Drill and practice are included in this category. "Discussion/Random" as used in this instrument is a very technical term. It is defined on the basis of our experience as "talk, talk, talk," or irrelevant verbal behavior as contrasted with logically structured discourse and represents the kind of stream-of-consciousness talk that seems to have little relevance to the purposes of the lesson. When discussion reveals a more purposeful direction it is classified as "Logical Thinking." Any time during which the teacher encourages thinking to take place, permits it to take place, or participates in the thinking process is listed as Logical Thinking. "The Thinking Process" refers to the time in which the teacher consciously points out ways of improving the thinking process, clarifies the problems, gets students to see the sources of their difficulty, helps supply the necessary information or the necessary technique for attacking a problem, or drives home the process through which students arrive at conclusions or make judgments.

---

* See Appendix B for instructions for use of TCAP and definitions and illustrations of terms.

Both types of Management represent first-level intellectual activity.* Presentation and Recitation/Drill represent second-level. Discussion/Random may fall into either level one or level two. Logical Thinking represents the third level and Attention to the Thinking Process represents the fourth level.

It should be emphasized that no attempt was made to judge the student teacher's behavior as to quality other than to place it in a category. Imperative or affective statements were not judged as accepting or rebuking. Presentation was judged neither for its accuracy nor its effectiveness. Logical Thinking and Attention to the Thinking Processes were not rated as to their rigor or precision.

The categories were determined on the basis of the way in which the observer, by looking at overt teacher behavior, viewed the teacher's intentions. Whether or not the student teacher succeeded was of secondary importance. No presumption was made as to what constituted a desirable profile. The study attempted to discover what kinds of patterns of teacher behavior obtained in classes taught by student teachers.

It was discovered that completed profiles were effective devices for helping the student to analyze his teaching and reconstruct his purposes and behavior.

The relationships between TCAP categories and the analyses of the various levels of intellectual activity, kinds of statements, the teacher's intentions, and clue words are shown in Figure 4.1.

Using the Teacher Classroom Activity Profile

In practice, the observer records a continuous line starting at the beginning of class, moving through one of the seven categories. When an activity appears to have some elements of two categories a vertical line is used to indicate this relationship. For instance, if the teacher is making a presentation but asks a question which prompts the class to do some thinking, a line is dropped down into the thinking category. Similarly, there is considerable fluidity among the various categories.

In actual practice, the three-minute intervals marked off in the form are frequently subdivided into thirds or one-minute intervals. At the end of the hour a summary is tabulated and the percentage of time spent in each of the categories is calculated.

Originally, the TCAP was designed to be used for only ten-minute intervals during the observation. Its use proved to be so valuable that subsequently it has been used to cover the total observation. The recording instrument is printed on no-carbon-required (NCR) paper and after an observation the TCAP is used as the basis for a supervisor-student teacher conference. A copy of it is left with the student teacher, another copy retained in the file on the student teacher, and the third copy used for research purposes.**

---

*See pages 25, 26 and 27 for explanation of levels of intellectual activity.

**A copy of the Teacher Classroom Activity Profile form is included in Appendix B.

Since this investigation was concerned with studying what the teacher does, the TCAP, like the SSTPP, focused exclusively on teacher behavior. Obviously there is interaction between teacher and pupil but for the purpose of this study only teacher behavior was recorded.

The TCAP, like the recording of classroom behavior as done by Medley and Mitzel, by B. O. Smith, by the Flanders-Amidon group, and by Arno Bellack, reports rather than evaluates. The TCAP does give a gross map of the kinds of activity in which the teacher engages in the classroom. The use of anecdotal notes and illustrative material on the TCAP observation form makes possible a useful "reconstruction" of the teacher's activity. This is not to deny that observers have to make some judgments as to which category appropriately describes the teacher's behavior. But this kind of judgment is necessary in any kind of categorization. There is no attempt to evaluate the quality of the teacher's activity as is the case with SSTPP.

It was discovered that observers could learn the techniques for using the TCAP in a relatively short amount of time and with a minimum of training. As a result, the TCAP was widely used by other persons in the student teaching program--departmental supervisers, cooperating teachers, and even pupils in the classes.

Table 4.1 presents a summary of teacher classroom activity as reported in 1,097 observations made by the Division of Teaching staff and 264 observations made by departmental supervisors during the two-year period 1966-67 and 1967-68. Table 4.2 presents in graphic form the time distribution divided between academic-type classes and laboratory-type classes by staff members. Table 4.3 gives the mean percentages of time devoted to each category.

Analysis of TCAP Data

The most striking fact to emerge from the use of the TCAP is the small proportion of time devoted to Thinking and the Thinking Process as here defined. This fact helps explain the relatively low reliability observers obtained on Criteria VII and VIII which assessed the logical nature of classroom discourse and conscious attention to the process of thinking. When only six minutes, or ten percent, of an hour's observation are providing data with which to assess the student teacher's performance one should not expect reliable results because of the limited time sample. This holds true also to a greater extent for Criterion VIII--Thinking Process which occupied only two percent of the time. Similarly, in the regression analysis Criteria VII and VIII played a very minor role in the global evaluation of the student. It seems quite logical that one of the reasons for this was the fact that only a limited amount of data was available to the observer on which to make his assessments.

An analysis of variance using the 2 x 2 factorial design was performed on each of the seven categories. One variable consisted of reports by Departmental Supervisors and reports by Staff. The other variable consisted of laboratory-type classes and academic-type classes. The means and the F ratio for these variables are reported in Table 4.4. In TCAP reports made by Division of Teaching Supervisors and Departmental Supervisors no significant differences were apparent except with respect to the percentage of time classified as Discussion, where the F

FIGURE 4.1

CUE CHART FOR CLASSIFYING TEACHER BEHAVIOR ON TCAP

TCAP CATEGORIES	LEVELS OF THINKING	KINDS OF STATEMENTS	TEACHER'S PRIMARY INTENTIONS OR CENTRAL PURPOSES	CLUE WORDS
Management--Non-Learning	Level One	Imperative--Affective	To control pupils To direct their behavior	Please (All commands) Class
Management--Learning (Originally included in Management)	Level One	Imperative--Affective	To influence classroom climate	Please (All commands)
Presentation	Level Two	Reporting--Stating--Recalling information from memory	To inform students	Now! Give me your attention Take notes on this
Recitation/Drill	Level Two	Questions and responses for getting students to report, state, recall information	To discover what pupils know or do not know To inform pupils	What - - - - ? How - - - - ?
Discussion/Random (Originally called Discussion I)	Levels One and Two	Affective--Responding without thinking--Stream-of-consciousness talk	To kill time To impress students as contrasted with informing them To establish relationships	One time I - - - - Did you hear about - - ? Ah! Ah! - - - -
Logical Thinking (Originally called Discussion II)	Level Three	Questions and statements which help students to think, to summarize, to reason, to evaluate, to analyze, to discover	To get students to think To guide students to discovery and verification	Why? followed by pause How? followed by pause What is the evidence? What are the assumptions? What are the inferences?
Attention to Thinking Process	Level Four	Statements about language and logic which help students to analyze the ways in which they come to empirical conclusions and normative judgments--Statements which help them locate sources of misunderstanding and disagreement	To provide students with the skills for thinking To inform them of accepted ways of using language To provide skills for pursuing truth	What do you mean? How do you define that? Is the sampling adequate? Where is our hang-up?

TABLE 4.1

MEAN PERCENTAGES OF TIME DEVOTED TO EACH CATEGORY OF TEACHER CLASSROOM ACTIVITY
AS REPORTED BY DEPARTMENTAL SUPERVISORS AND DIVISION OF TEACHING SUPERVISORS--FOR BOTH ACADEMIC AND LABORATORY CLASSES
Departmental Supervisors N = 264 Division of Teaching N = 1097

TCAP Categories	Percent of Time Observed 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	F Ratio	Level of Significance
Management--NL			
Supervisors	4.45	.19	
Staff	5.10		
Management--L			
Supervisors	22.99	.29	
Staff	27.03		
Presentation			
Supervisors	21.17	1.27	
Staff	20.44		
Recitation/Drill			
Supervisors	24.10	3.65	
Staff	26.33		
Discussion/Random			
Supervisors	16.71	7.81	P < .01
Staff	9.82		
Logical Thinking			
Supervisors	9.01	.03	
Staff	9.49		
Thinking Process			
Supervisors	1.48	.26	

TABLE 4.2

MEAN PERCENTAGES OF TIME DEVOTED TO EACH CATEGORY OF TEACHER CLASSROOM ACTIVITY FOR LABORATORY-TYPE AND ACADEMIC-TYPE CLASSES AS REPORTED BY DIVISION OF TEACHING STAFF

ACADEMIC N = 584 LABORATORY N = 513

TCAP Categories	Percent of Time Observed (0–30)	F Ratio	Level of Significance
Management--NL			
Academic	3.74	11.56	.005
Laboratory	6.66		
Management--L			
Academic	27.25	1.20	
Laboratory	26.76		
Presentation			
Academic	23.32	15.84	.005
Laboratory	17.17		
Recitation/Drill			
Academic	23.75	4.04	.05
Laboratory	29.27		
Discussion/Random			
Academic	6.27	27.37	.005
Laboratory	13.85		
Logical Thinking			
Academic	12.89	9.67	.005
Laboratory	5.62		
Thinking Process			
Academic	2.75	8.00	.005
Laboratory	.63		

TABLE 4.3

MEAN PERCENTAGES OF TIME DEVOTED TO EACH CATEGORY OF TEACHER CLASSROOM ACTIVITY BY OBSERVERS AND BY TYPE OF CLASS

TCAP Category	Departmental Observers		Staff Observers		All Observers and All Classes
	Lab N = 112	Academ. N = 152	Lab N = 513	Academ. N = 584	
Management--NL	5.84	3.42	6.66	3.74	5.10
Management--L	22.16	23.60	26.76	27.25	27.03
Presentation	15.98	25.65	17.17	23.32	20.44
Recitation/Drill	23.59	24.47	29.27	23.75	26.33
Discussion/Random	22.34	12.55	13.89	6.27	9.82
Logical Thinking	9.84	8.40	5.62	12.89	9.49
Thinking Process	.93	1.88	.63	2.75	1.76

ratio of 7.81 showed a significant difference at the .01 level as reported in Table 4.1. Departmental Supervisors reported approximately 17 percent of classroom time devoted to Discussion/Random as contrasted with 10 percent reported by the Division of Teaching staff. The tendency for Departmental Supervisors, as contrasted with Division of Teaching Supervisors, to report a larger proportion of time spent in this type of activity obtains in both academic-type classes and laboratory-type classes. One explanation might be the highly technical meaning applied to Discussion/Random. The Division of Teaching Supervisors participated in the development of TCAP, the clarification of terms, training sessions, and trial runs. Inasmuch as the Departmental Supervisors had only limited training, it is likely that the more common connotation of the term "Discussion/Random" was used.

An analysis of variance made of TCAP reports on laboratory-type classes and academic-type classes revealed a significant difference in the distribution of class time. The differences were significant at the .005 level for all categories except Recitation where it was significant at the .05 level and Management--Learning where there was no significant difference.

As shown in Table 4.3, laboratory-type classes devote relatively more time to Management--NL and Discussion/Random than do academic-type classes. The data suggest, as one would expect, that it takes more time to organize equipment and distribute materials in laboratory classes than in academic classes. Music classes need to get instruments ready; physical education classes require special clothing; science laboratories need to arrange their equipment. It seems a bit surprising that the difference is not greater: 6.66 percent, or four minutes, of laboratory-class time as contrasted with 3.74 percent, or two minutes, of academic-class time.

Again as one would expect, the data suggest that teachers in academic classes spend more time in Presentation--23.3 percent, or twelve minutes--than do teachers in laboratory classes with 17.1 percent, or eight minutes. That the difference is not greater is due to the fact that when teachers talk to individuals or small groups in laboratory-type classes it is classified as Presentation.

The higher proportion of time spent in Recitation/Drill for laboratory-type classes was probably due to the inclusion, by definition, of drill and practice in this category.

The larger proportion of laboratory time spent in Discussion/Random probably reflects an accurate picture of irrelevant discussion and pointless talk.

The data suggest, again as one would expect, that greater opportunities to participate in the higher levels of thinking and analysis arise in the academic classes. The 12.89 percent, or six minutes, and the 5.62 percent, or three minutes, devoted to Logical Thinking in the respective types of classes are similarly disturbing because they represent such a small fraction of the total class time.

TABLE 4.4

ANALYSIS OF VARIANCE ON DISTRIBUTION OF CLASSROOM TIME BY PERSON MAKING OBSERVATION AND BY TYPE OF CLASS

A = Laboratory Class--Academic Class
B = Departmental Supervisor--Staff Supervisor

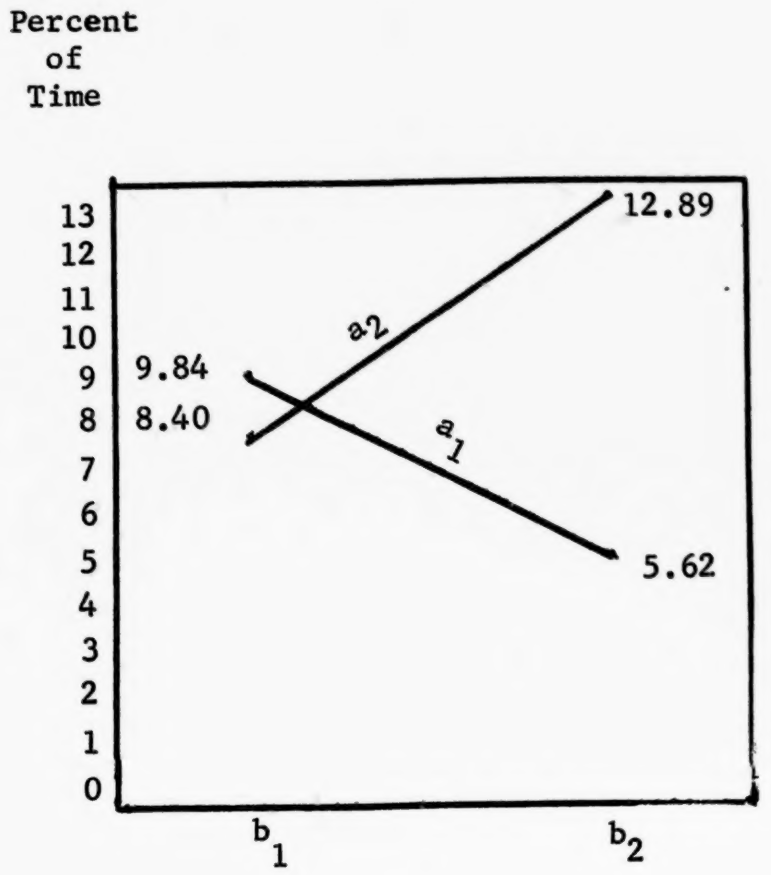
Time Category	Source of Variance	DF	MS	F
Management--NL	A	1	1.43	11.56***
	B	1	.02	.20
	AB	1	.00	.01
	Error	1047	.12	
Management--L	A	1	1.11	1.20
	B	1	.27	.29
	AB	1	.36	.39
	Error	1047	.92	
Presentation	A	1	15.85	22.30***
	B	1	.91	1.28
	AB	1	1.59	2.23
	Error	1047	.71	
Recitation/Drill	A	1	3.50	4.04*
	B	1	3.17	3.66
	AB	1	.80	.92
	Error	1047	.87	
Discussion/Random	A	1	27.37	36.51***
	B	1	5.86	7.81**
	AB	1	.01	.01
	Error	1047	.75	
Logical Thinking	A	1	3.06	9.68***
	B	1	.01	.04
	AB	1	3.84	12.15***
	Error	1047	.32	
Thinking Process	A	1	.325	8.01***
	B	1	.01	.28
	AB	1	.02	.45
	Error	1047	.04	

* F(1,1047;05) = 3.84
** F(1,1047;01) = 6.32
*** F(1,1047;005) = 7.88

An interesting relationship appears in the interaction analysis with respect to the Thinking category. When the academic and laboratory classes and the Departmental and Staff classifications are used, as reported in Table 4.4, only Logical Thinking shows any significance but does so at the .005 level. Figure 4.2 presents in graphic form the interaction effect with respect to Logical Thinking.

FIGURE 4.2

GEOMETRIC PROFILE OF INTERACTION BETWEEN TYPES OF OBSERVERS AND TYPES OF CLASSES ON LOGICAL THINKING CATEGORY



A Types of Class
a_1 Laboratory
a_2 Academic

B Types of Observers
b_1 Departmental
b_2 Staff

Apparently Departmental Supervisors and Staff Supervisors do not perceive the Thinking category the same in laboratory- and academic-type classes. Departmental Supervisors tend to see more Thinking in laboratory classes than in academic classes whereas the reverse is true for Staff Supervisors.

One explanation for this variance could be that Departmental and Staff Observers interpreted the definition of Thinking differently. For laboratory-type classes time in which the teacher is actively involved in helping the students as they pursue their own purposeful activity is classified as Thinking.

Another possible explanation could be that the subject-matter specialists who are the Departmental Observers tend to identify more Thinking in laboratory-type activity because of their special subject-matter orientation while on the other hand Staff Observers, who are generalists, are more comfortable in the academic classroom than in the more specialized laboratory. This is to say that a Music specialist in observing an orchestra rehearse tends to consider more rehearsal time as belonging to the Thinking category. A Physical Education specialist considers the organized game as a Thinking experience or the Home Economics supervisor sees considerable Thinking as the girls bake the cake.

The staff generalists apparently find it easier to identify Thinking in the academic classes because they are more familiar with this type of class. If it were possible to pursue this question further some light might be cast upon the relative merits of supervision by generalists as opposed to subject-matter specialists.

Examination of Table 4.5 shows the ways in which time distributions vary with subject fields. It is interesting to note that the similarities far exceed the differences. While some variations do occur, the main features of the profile present the same pattern in all subject fields.

Art teachers spent considerably less time than the average teachers in Presentation and Logical Thinking and more time in Discussion/Random. Teachers of Business devoted more than average amounts of time to Management--Learning and Discussion/Random but less to Presentation and Logical Thinking. Home Economics teachers differ from the average by spending more time in Presentation and Logical Thinking and less time in Recitation/Drill. Industrial Arts varies from the pattern primarily by devoting less time to Recitation/Drill and Logical Thinking and more time to Discussion/Random. Music teachers spend more time than average in Management--Non-Learning and almost twice as much in Recitation/Drill, with less time in Discussion/Random and Logical Thinking. Physical Education teachers approximate the pattern except for a larger percentage of time in Management--Non-Learning and Recitation/Drill and a smaller percentage in Logical Thinking and the Thinking Process. English teachers devote more than average time to Management--Learning, Logical Thinking, and the Thinking Process and less time to Recitation/Drill and Discussion/Random. Foreign Language teachers differ from the average by spending less time in both Management categories and Presentation (26 percent) and much more time in Recitation/Drill (44 percent). Social Studies teachers contrast with the average in greater time devoted to Presentation, Logical Thinking, and the Thinking Process. Less time is spent in Recitation/

TABLE 4.5

MEAN PERCENTAGES OF TIME DEVOTED TO THE VARIOUS CATEGORIES OF TEACHER CLASSROOM ACTIVITY BY SUBJECT FIELDS AND TOTAL FOR ALL FIELDS

TCAP Categories	Total	Art	Bus.	Eng.	For. Lang.	Home Ec.	Ind. Arts	Math	Music	Phys. Ed.	Sci.	Soc. St.	Sp. Ed.
	N=1097	N=37	N=97	N=170	N=73	N=58	N=101	N=86	N=85	N=135	N=65	N=154	N=36
Management--NL	5.10	5.45	4.24	4.45	3.56	3.51	6.34	2.69	9.37	8.60	2.56	4.10	3.83
Management--L	27.03	23.54	34.09	33.14	20.76	26.67	26.98	25.98	15.84	29.15	29.10	25.30	20.69
Presentation	20.44	10.08	13.69	19.43	13.94	28.74	18.88	25.86	14.29	17.19	31.32	31.31	6.02
Recitation/Drill	26.33	24.24	26.29	18.80	44.05	16.44	20.05	19.62	48.21	33.28	20.09	16.87	51.83
Discussion/Random	9.82	31.97	14.84	7.62	7.20	11.00	21.63	5.61	6.87	7.98	3.98	5.20	8.36
Logical Thinking	9.49	4.45	5.75	14.45	8.91	11.82	5.87	16.41	4.89	3.48	11.12	13.46	5.86
Thinking Process	1.76	0.24	1.07	2.26	1.53	1.79	0.24	3.69	0.52	0.29	1.75	3.58	3.47

Drill and Discussion/Random. Mathematics teachers show the highest percentage of time of any subject in Logical Thinking and Thinking Process with less time spent in Management--Learning, Recitation/Drill, and Discussion/Random. Science teachers spend less time in Management--Learning, Recitation/Drill, and Discussion/Random but considerably more on Presentation. Special Education teachers differ from the average primarily in the small amount of time devoted to Presentation and the large amount (52 percent) devoted to Recitation/Drill.

The patterns of student teacher classroom activity were remarkably stable over time. Since the TCAP was discussed in considerable detail with the student teacher one could expect that there would be considerable change in the way the student spent his time between the first and second observations. From two to three weeks elapsed between the two observations. Inasmuch as both students and supervisors expressed concern about the small proportion of time devoted to Logical Thinking and the Thinking Process, one would expect a marked increase in time spent in these activities. Table 4.6 reveals that the patterns of teacher activity, on the average, were almost identical in the first and second observations.

TABLE 4.6

MEAN PERCENTAGES OF TIME SPENT IN VARIOUS TYPES OF CLASSROOM ACTIVITY DURING FIRST AND SECOND OBSERVATIONS

N = 563

Category	First Observation	Second Observation
Management--Non-Learning	4.71	4.11
Management--Learning	23.52	21.31
Presentation	15.55	16.30
Recitation/Drill	20.56	20.78
Discussion/Random	7.90	7.14
Logical Thinking	7.39	7.78
Thinking Process	1.20	1.42

These data would tend to support Marie Hughes' conclusion that teachers' behavior patterns are stable through time.[1]

---

[1] Marie Hughes, Development of the Means for Assessment of the Quality of Teaching in Elementary Schools (Salt Lake City: University of Utah Press, 1959).

Comparison with Other Research

No comparable studies of teacher classroom behavior were found. B. O. Smith's preliminary report on his study of logic summarizes the proportionate number of entries classified in each of his 13 categories for 17 different teachers for five consecutive class periods per teacher or a total of 85 classes, as shown in the following table.[2]

TABLE 4.7

PROPORTION OF LOGICAL ENTRIES REPORTED BY SMITH

Category	Percent of Total Number of Entries
Defining	4.1
Describing	25.3
Designating	14.8
Stating	6.8
Reporting	2.9
Substituting	.3
Valuating	4.6
Opining	5.3
Classifying	3.0
Comparing and Contrasting	3.3
Conditional Inferring	7.3
Explaining	12.9
Directing and Managing Classroom	9.4

While Smith's report does not purport to be a time study, the fact that only approximately 10 percent of the entries were classified as management and 90 percent as some type of logical operation makes one realize that there must be vast differences either in the classroom situations examined in the two studies or in the means of observing and/or reporting. It is recognized that Smith combined the teacher behaviors and students' behaviors while this study reported only teacher activity.

This study found that approximately 25 percent of the time, as contrasted with 10 percent of the entries in the Smith study, was spent in managing the classroom. Approximately an additional 15 percent of the time was spent in Discussion/Random--a category which apparently has no counterpart in the Smith study.

Smith's first four categories, which account for 53.9 percent of the entries, would likely be classified under Presentation and/or Recitation, which account for approximately 45 percent of the classroom time.

---

[2] B. O. Smith, A Study of the Logic of Teaching. USOE Project No. 258 (7257), Mimeographed (University of Illinois: Urbana, 1959).

Smith's next seven categories, which account for 36.7 percent of his entries, would be classified as Logical Thinking and the Thinking Process, which together consumed only 11 percent of the classroom time in this study.

Regardless of whether or not these data are subject to any type of comparison, the staff is deeply concerned by the relatively small amount of time devoted to the third and fourth levels of intellectual activity.

Marie Hughes suggests that the following range of percentages in each of the seven behavior categories produced the optimum interaction for learning in the elementary school:[3]

Controlling Functions	20-40%
Imposition	1-3 %
Facilitating	5-15%
Content Development	20-40%
Personal Response	8-20%
Positive Affectivity	10-20%
Negative Affectivity	3-10%

While this study was concerned only with secondary student teachers and the categories are not identical, it is interesting to see how the results might be compared with the time divisions recommended by Hughes, as shown in Table 4.8, which follows.

TABLE 4.8

COMPARISON OF PERCENTAGES OF TIME SPENT IN TCAP CATEGORIES AND HUGHES' CATEGORIES

TCAP Categories	% of Time	Hughes' Categories	% of Time
Management--NL & L	27	Controlling Imposition	20-40 1-3
Presentation and Recitation/Drill	45	Content Development	20-40
Discussion/Random	16	Personal Response Positive Affectivity Negative Affectivity	8-20 10-20 3-10
Logical Thinking	10	(Included in Content Development)	
Thinking Process	2	Facilitating	5-15

---

[3] Hughes, op. cit.

Conclusions

The Teacher Classroom Activity Profile has proven to be an effective instrument in the supervision of student teachers. Its use has enabled the supervisor and the student teacher to approach their conference on a more objective basis, looking at what the student teacher did rather than evaluating how well he did it. The evaluation is sometimes entirely unnecessary or, rather, is self-evident once the supervisor and the student teacher agree on what the student teacher actually did. The use of the TCAP has also resulted in a shift in the pre-student-teaching orientation. Students are apprised of it and the importance attached to the last two categories is made clear. As a result, student teachers are beginning to plan their work so as to pay more attention to thinking. The questions they ask are of a higher order and recent observations suggest a marked increase in the proportion of time devoted to Logical Thinking and the Process of Thinking.

The fact that student teachers devoted only ten percent of their class time to what is considered Thinking and only one percent to the Thinking Process has serious implications for their educators. This study did not examine certificated teachers. However, there is a suggestion that the differences in classroom behavior are not very great between student teachers and supervising teachers. Since supervising teachers, in the main, have not expressed concern about the way student teachers distribute their classroom activity, one might conclude that they are not disturbed by the small amount of time devoted to Thinking and the Thinking Process.

This study does not presume to establish an ideal model of classroom behavior. Perhaps spending ten percent of class time in what has been defined as Thinking is all that can be expected. However, the literature of education either explicitly or implicitly suggests that the chief function of the schools in a democracy is to improve the intellectual ability of the people. The data presented here suggest that student teachers either reject this purpose in favor of doing other things or are not prepared to help students learn to think, or the concept of Thinking and the Thinking Process as formulated for this study needs further refinement before it becomes a meaningful construct. The staff is convinced that teacher educators need to come to grips with this problem even though it presents great difficulties.

CHAPTER V

ANALYSIS OF THE SECONDARY STUDENT TEACHER PERFORMANCE PROFILE (SSTPP)

Beginning in the fall of 1966, all staff members of the Division of Teaching used the Secondary Student Teacher Performance Profile (SSTPP) together with the Guide as a regular part of the supervisory process. The staff also used it under five experimental conditions as reported in Chapter III. The data examined in this chapter are drawn from a three-year period. Since the other data with which comparisons were made were not universally available, special care has been taken to report the number of students involved in each analysis.

Three sources of data were used. The first was that available from the regular university records--cumulative grade-point average, scores on the Scholastic Aptitude Test taken as a freshman, and estimated grade-point average computed from SAT scores and rank in high school class. A second source consisted of examinations administered as a part of this research--Cattell's The Sixteen Personality Factor Questionnaire, Form A and the Cornell Test of Critical Thinking, Form Z, by Robert Ennis.* The third source of data was reports by university personnel--assessments made by Division of Teaching staff in preliminary interviews, assessments made by Departmental Supervisors in preliminary interviews, and evaluations made by the cooperating teachers at the end of the student teaching period.

The experimental group consisted of 564 individual secondary student teachers selected at random from those assigned during the academic years 1966-67 and 1967-68. The total group included, in addition to the experimental group, 243 student teachers from the years 1966-68 and 264 from the year 1965-66.

The Observers

Figure 5.1 presents in graphic form a summary of criterion scores assigned by the twelve different staff observers who worked with the project at one time or another, in 2,018 classroom observations.

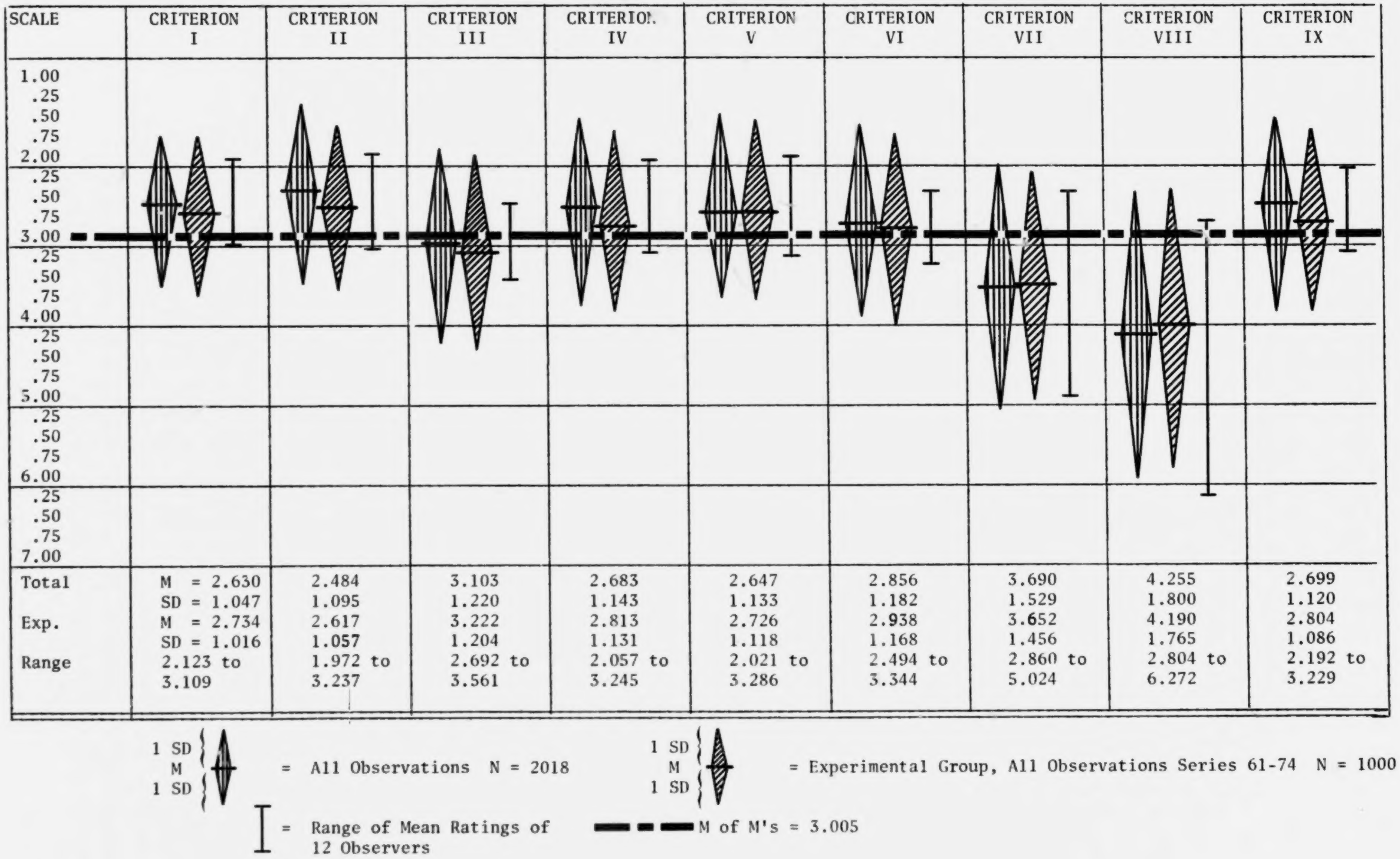
Figure 5.1 shows, for example, on Criterion I (Understanding, Friendly vs. Egocentric, Aloof) the mean estimate given all student teachers observed by the twelve staff members was 2.63, with a standard deviation of 1.05. The mean estimate in 1000 observations of the 564 students in the experimental group was 2.73 with a standard deviation of 1.02. The mean

---

*Plans to administer the National Teacher Examination were abandoned when students failed to volunteer to take the examination even though it was offered at no financial cost to them. The students were forced to take the 16 PF test and the Cornell test. It is possible that their scores on these tests could have been distorted by their attitude toward being required to take them.

FIGURE 5.1

MEANS, STANDARD DEVIATIONS, AND RANGES OF CRITERION ESTIMATES
MADE BY TWELVE OBSERVERS FOR TOTAL GROUP AND EXPERIMENTAL GROUP



SCALE	CRITERION I	CRITERION II	CRITERION III	CRITERION IV	CRITERION V	CRITERION VI	CRITERION VII	CRITERION VIII	CRITERION IX
Total	M = 2.630	2.484	3.103	2.683	2.647	2.856	3.690	4.255	2.699
	SD = 1.047	1.095	1.220	1.143	1.133	1.182	1.529	1.800	1.120
Exp.	M = 2.734	2.617	3.222	2.813	2.726	2.938	3.652	4.190	2.804
	SD = 1.016	1.057	1.204	1.131	1.118	1.168	1.456	1.765	1.086
Range	2.123 to 3.109	1.972 to 3.237	2.692 to 3.561	2.057 to 3.245	2.021 to 3.286	2.494 to 3.344	2.860 to 5.024	2.804 to 6.272	2.192 to 3.229

estimates made by all twelve staff members on Criterion I ranged from 2.12 to 3.11. The mean estimate assigned to all criteria on all observations was 3.01. It should be kept in mind that the most favorable rating was 1 and the least favorable 7.

The mean estimates and standard deviations are similar for the total group and for the experimental group.

It is apparent that the cultural bias toward generosity of ratings so prevalent in America moved the midpoint approximately one unit higher than the midpoint of a seven-point scale.* This J effect applies to all criteria except VII and VIII, which are concerned with logic.

Figure 5.1 illustrates by a different treatment of the data several conclusions reported in Chapter III. Staff supervisors tended to agree most closely with respect to Criteria III (Stimulating, Imaginative, Surgent vs. Dull, Routine, Unimaginative) and VI (Communicates Well and Empathetically vs. Communicates Ineffectively and Perfunctorily). They tended to vary the most on Criteria VII (Classroom Discourse Characterized by Reasoning and Creative Thinking vs. Classroom Discourse Characterized by Simple Recall) and VIII (Directs Attention to Process of Thinking vs. Fails to Call Attention to Process of Thinking). They tended to rate student teachers most favorably on Criterion II (Planned, Organized, Responsible vs. Unplanned, Disorganized, Irresponsible). Similar favorable ratings were given on Criteria I (Understanding, Friendly vs. Egocentric, Aloof), IV (Perceives Self as Competent vs. Perceives Self as Less Than Adequate), V (Has Mastery of Facts and Organizing Principles of Field vs. Has Only Minimum Knowledge of Field) and IX (High Professional Potential vs. Low Professional Potential). Less favorable ratings were given on Criteria VI (Communicates Well and Empathetically vs. Communicates Ineffectively and Perfunctorily) and III (Stimulating, Imaginative, Surgent vs. Dull, Routine, Unimaginative). Student teachers were rated most unfavorably on Criterion VIII (Directs Attention to Process of Thinking vs. Fails to Call Attention to Process of Thinking) and Criterion VII (Classroom Discourse Characterized by Reasoning or Creative Thinking vs. Classroom Discourse Characterized by Simple Recall).

---

* The Stanford Teacher Competence Appraisal Guide which used a seven-point scale was arranged to reduce the J-curve effect and overcome the generosity factor by designating 3 as the average. It also indicated a percentage figure and a name for each point on the scale, as follows: 1 = Weak, 30%; 2 = Below Average, 15%; 3 = Average, 15%; 4 = Strong, 15%; 5 = Superior, 15%; 6 = Outstanding, and 7 = Truly Exceptional, 10%. No percentages were given and no names were assigned to the seven points of the scale in this study. Only the two extremes were identified.

R. N. Bush, _et al._, _op. cit._, p. 5.

TABLE 5.1

MEAN ESTIMATES ON EACH CRITERION OF SSTPP TOGETHER WITH STANDARD DEVIATIONS BY OBSERVERS

	I	II	III	IV	V	VI	VII	VIII	IX
Observer #12 N = 290									
M	2.462	2.451	3.020	2.683	2.629	2.648	3.486	3.880	2.858
SD	0.980	1.025	1.190	0.966	0.896	0.995	1.672	1.804	0.998
Observer #13 N = 246									
M	3.069	2.418	3.317	2.885	2.857	3.280	3.884	5.375	2.703
SD	1.191	1.327	1.351	1.248	1.336	1.238	1.592	1.589	1.302
Observer #14 N = 194									
M	2.123	2.444	2.769	2.550	2.594	2.494	5.024	6.272	2.505
SD	0.952	1.119	1.215	1.369	1.241	1.316	1.508	1.176	1.191
Observer #15 N = 208									
M	2.322	2.163	2.692	2.057	2.304	2.529	3.524	3.712	2.298
SD	0.925	0.954	1.077	0.966	0.984	1.093	1.249	1.422	0.941
Observer #16 N = 328									
M	3.109	2.926	3.505	2.913	3.035	3.128	3.988	5.191	3.064
SD	0.981	0.979	1.125	1.103	1.060	1.152	1.540	1.864	1.011
Observer #17 N = 192									
M	2.421	2.131	2.989	2.718	2.021	2.651	3.097	3.910	2.551
SD	0.846	0.822	1.015	0.978	0.743	0.974	1.237	1.621	0.809
Observer #18 N = 182									
M	2.412	1.972	2.967	2.329	2.626	2.835	3.443	3.772	2.192
SD	1.248	1.122	1.478	1.175	1.434	1.424	1.693	1.842	1.367

TABLE 5.1

MEAN ESTIMATES ON EACH CRITERION OF SSTPP TOGETHER WITH STANDARD DEVIATIONS BY OBSERVERS

(Continued)

	I	II	III	IV	V	VI	VII	VIII	IX
Observer #19 N = 122									
M	3.098	3.237	3.561	3.245	3.286	3.344	3.909	4.105	3.229
SD	0.720	0.953	1.023	1.108	1.040	1.010	1.229	1.198	1.066
Observer #23 N = 58									
M	2.896	2.666	3.105	2.810	2.517	2.879	3.022	3.021	2.827
SD	0.831	1.023	0.976	0.887	0.903	1.044	0.965	0.966	0.819
Observer #24 N = 97									
M	2.175	2.484	2.812	2.474	2.257	2.708	3.452	3.177	2.721
SD	0.595	0.817	1.039	0.902	0.807	1.004	0.949	1.055	0.838
Observer #25 N = 51									
M	2.372	2.450	2.941	2.901	2.400	2.862	2.860	2.804	2.470
SD	1.019	1.101	1.377	1.330	1.142	1.265	1.337	1.360	1.254
Observer #26 N = 50									
M	2.780	2.700	3.280	2.920	2.687	2.680	3.400	3.280	2.780
SD	0.581	0.814	0.701	0.665	0.657	0.740	0.808	0.833	0.763

Table 5.1 presents the mean estimates on each criterion by observer. For example, it shows that Observer 12 reported 290 observations with an average rating on Criterion I of 2.46 and a standard deviation of .98.

While it is possible that the students observed by different members of the staff did vary on the criteria it is likely that personal bias of observers accounted for much of the variation shown. This bias seems to persist in spite of intensive observation training sessions. For example, Observers 12, 14, 15, and 24 tended to rate students higher on Criterion I (Understanding, Friendly vs. Egocentric, Aloof) than did Observers 13, 16, and 19. Observer 19 tended to rate students lower on all the criteria relating to personality than did other observers. Observers 13 and 14 tended to rate students lower on the criterion related to logic.

Brown examined this problem and concluded: "In order to understand what any given observer-judge means when he says, 'This is a good teacher,' or, 'This is a poor teacher,' we must know who the observer-judge is, what he values, and where he stands on important educational issues when he observes and judges."[1]

This study was designed to handle the problem of observer bias in two ways--first, by randomization of assignment and, second, by training. No provisions were made to isolate and correct for observer bias.

A preliminary investigation using factor analysis suggested that it would be possible to identify observer bias and determine a correction index. Such an investigation was beyond the purview of this study.

Platoons and Semesters

A special analysis was made of criterion ratings to see if any variation occurred between first and second platoon student teachers or between semesters. All ratings of "N" (not sufficient evidence) were deleted from the analysis. Figure 5.2 presents in graphic and tabular form the means and standard deviations of ratings on the nine original criteria and the three composite criteria. Table 5.2 presents the analysis of variance due to platoons and semesters.

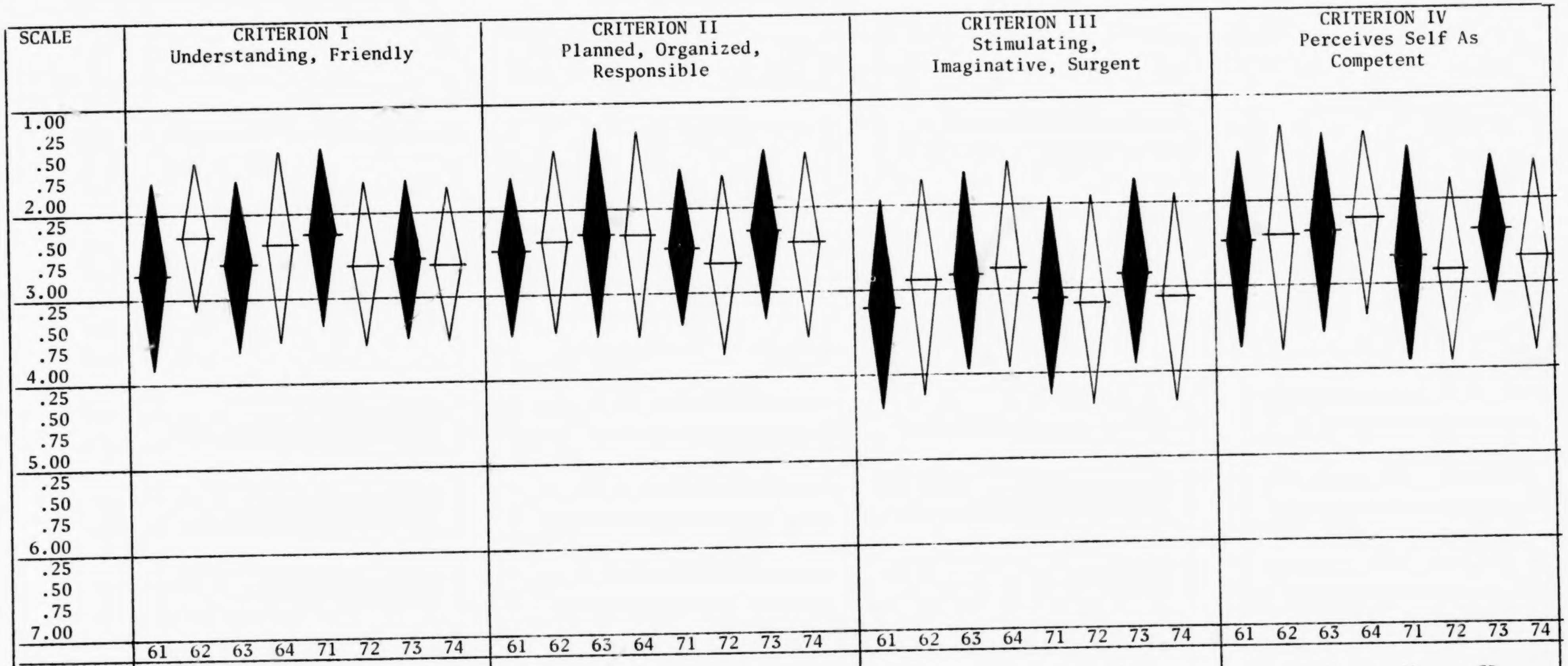
No significant variation occurred between the two platoons on any of the criteria. Inasmuch as Platoon I students have not taken the related Education courses nor Special Methods courses prior to student teaching while Platoon II students have completed these courses prior to their assignment, one might expect some differences in cri-

---

[1] Bob Burton Brown, "Observer-Judge Ratings of Teacher Competence," Childhood Education, Vol. 44 (November, 1967), p. 205.

FIGURE 5.2

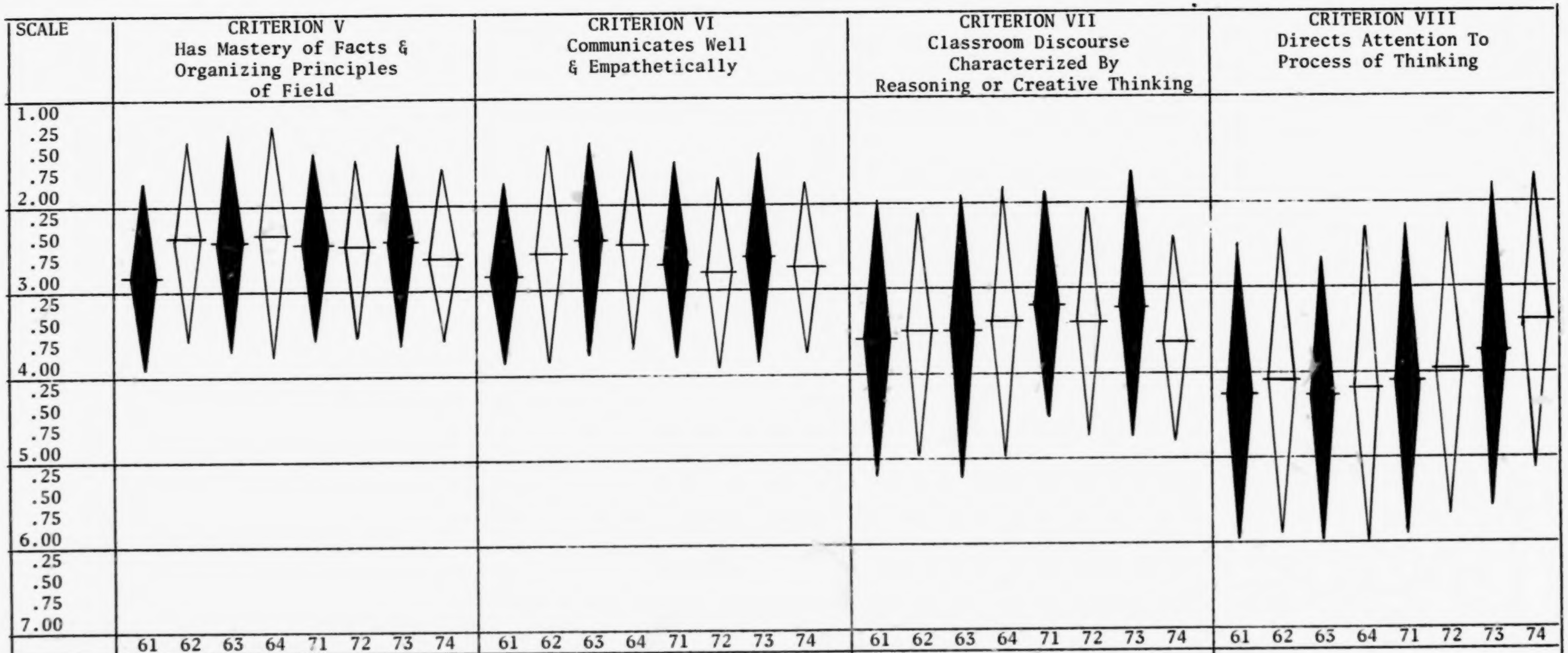
DISTRIBUTIONS OF MEANS AND STANDARD DEVIATIONS OF ESTIMATES
ON SECONDARY STUDENT TEACHER PERFORMANCE PROFILE, SECOND OBSERVATION, BY PLATOONS



PLATOON	N	M	SD	N	M	SD	N	M	SD	N	M	SD
61	60	2.85	1.08	60	2.68	0.96	60	3.36	1.20	60	2.68	1.11
62	55	2.41	0.87	55	2.50	1.10	55	3.09	1.30	55	2.56	1.30
63	78	2.74	1.02	78	2.43	1.23	76	2.96	1.22	78	2.53	1.19
64	74	2.50	1.16	73	2.45	1.20	74	2.89	1.23	74	2.39	1.10
71	59	2.40	1.08	57	2.61	0.92	57	3.28	1.16	58	2.77	1.25
72	67	2.74	0.98	66	2.78	1.01	67	3.32	1.21	67	2.97	1.07
73	90	2.70	0.94	90	2.41	0.98	90	2.98	1.09	90	2.50	0.87
74	81	2.75	0.92	81	2.59	1.12	81	3.29	1.20	80	2.81	1.09

FIGURE 5.2

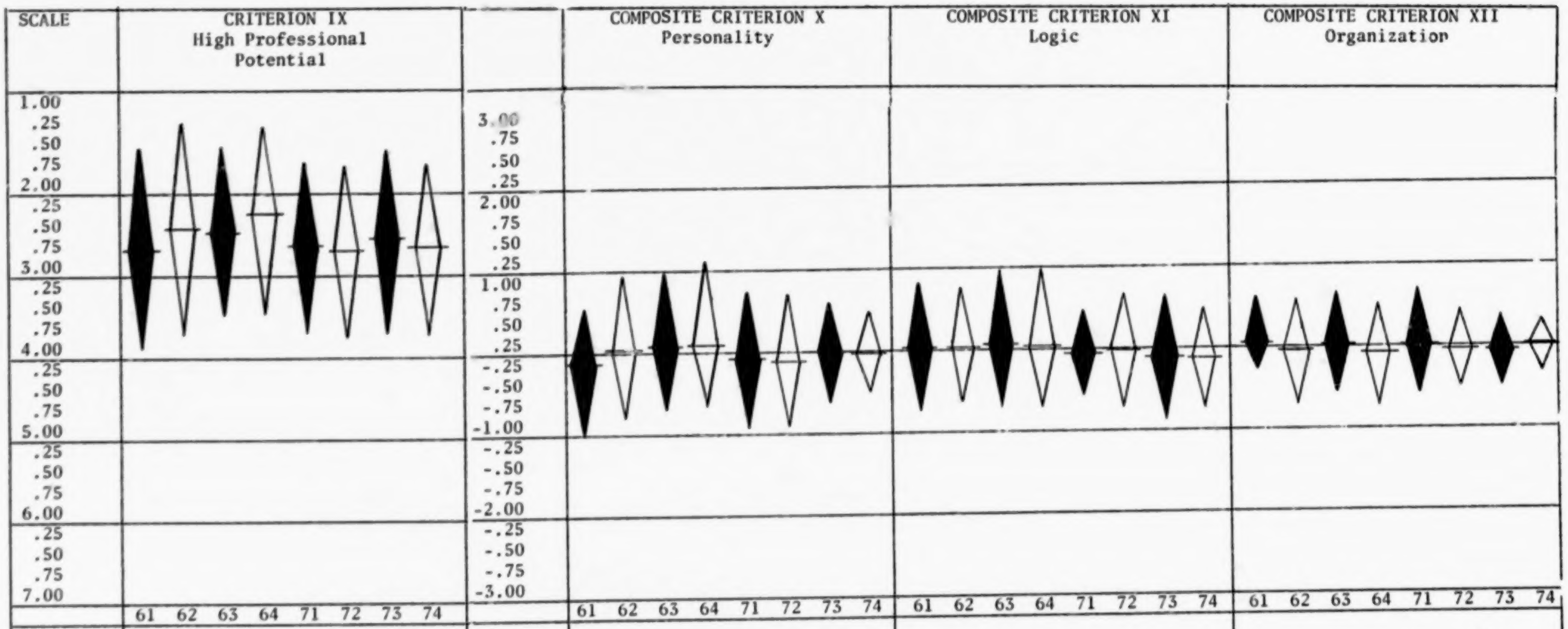
DISTRIBUTIONS OF MEANS AND STANDARD DEVIATIONS OF ESTIMATES
ON SECONDARY STUDENT TEACHER PERFORMANCE PROFILE, SECOND OBSERVATION, BY PLATOONS
(Continued)



PLATOON	N	M	SD	N	M	SD	N	M	SD	N	M	SD
61	60	2.95	1.11	59	2.91	1.07	43	3.74	1.57	32	4.40	1.72
62	55	2.54	1.18	54	2.72	1.26	45	3.68	1.44	38	4.23	1.77
63	77	2.59	1.19	78	2.62	1.22	66	3.69	1.61	56	4.46	1.69
64	73	2.52	1.32	72	2.66	1.16	58	3.51	1.60	49	4.30	1.85
71	59	2.62	1.09	58	2.81	1.14	44	3.31	1.30	42	4.23	1.81
72	67	2.64	1.06	66	2.93	1.13	52	3.53	1.32	48	4.10	1.69
73	89	2.58	1.14	90	2.74	1.23	73	3.34	1.59	74	3.82	1.90
74	80	2.76	0.93	81	2.86	0.98	68	3.75	1.21	58	3.53	1.74

FIGURE 5.2

DISTRIBUTIONS OF MEANS AND STANDARD DEVIATIONS OF ESTIMATES
ON SECONDARY STUDENT TEACHER PERFORMANCE PROFILE, SECOND OBSERVATION, BY PLATOONS
(Continued)



PLATOON	N	M	SD
61	60	2.81	1.18
62	55	2.56	1.27
63	77	2.61	1.04
64	74	2.44	1.18
71	58	2.79	1.03
72	67	2.85	1.04
73	90	2.71	1.11
74	81	2.82	1.02

PLATOON	N	M	SD	N	M	SD	N	M	SD
61	31	-0.25	0.90	31	0.06	0.90	31	0.17	0.58
62	37	0.05	1.00	37	0.07	0.81	37	-0.04	0.76
63	55	0.13	0.96	55	0.14	0.97	55	0.05	0.73
64	49	0.22	1.00	49	0.12	0.94	49	-0.07	0.75
71	38	-0.11	0.94	38	-0.06	0.64	38	0.06	0.77
72	45	-0.14	0.90	45	-0.01	0.81	45	-0.03	0.58
73	73	-0.00	0.74	73	-0.11	0.89	73	-0.06	0.57
74	57	-0.02	0.60	57	-0.13	0.72	57	0.01	0.45

terion estimates. However, this finding is consistent with other studies which have shown no significant differences between the two platoons.[2]

A slight variance, significant at the .05 level, occurred with respect to Criterion VIII by semesters. Examination of Figure 5.2 shows a step-wise improvement in the ratings assigned on Criterion VIII over the last three semesters. One explanation could be that student teachers became aware of the value attached to this criterion even though no special effort was made to instruct them. Another explanation could be that the Supervisors through experience and in-service training sessions, became more expert in identifying classroom behavior which directed pupils' attention to the thinking process.

The tabular information presented below Figure 5.2 shows that the most "N" (not-sufficient data) responses were made on Criteria VII and VIII, even through the last semester. For example in platoon 74, which was the fourth platoon of the year 1967-68, 81 ratings were made on Criteria I, II, III, VI, and IX, while only 68 ratings were made on Criterion VII, 58 on Criterion VIII, and one "N" was marked on Criterion V.

The number of subjects included in the composite criteria was substantially lower than the total students studied because only those students who had been rated on all criteria were included.

---

[2] Marvin A. Henry, The Relationship of Difficulties of Teachers to Selected Aspects of the Professional Sequence of Education (Unpublished doctoral dissertation. Bloomington, Indiana: Indiana University, June, 1963).

TABLE 5.2

ANALYSIS OF VARIANCE OF CRITERION RATINGS BY SEMESTER AND BY PLATOON

A = Semester
B = Platoon

Criterion	Source of Variance	DF	MS	F
I. Understanding, Friendly	A	3	.54	.50
	B	1	.98	.92
	AB	3	2.19	2.05
	Error	192	1.07	
II. Planned, Organized	A	3	1.77	1.61
	B	1	0.00	0.00
	AB	3	.51	.46
	Error	192	1.10	
III. Stimulating	A	3	3.25	1.98
	B	1	1.00	.61
	AB	3	2.47	1.51
	Error	192	1.64	
IV. Self Concept	A	3	2.34	1.73
	B	1	.14	.10
	AB	3	1.21	.90
	Error	192	1.35	
V. Subject Mastery	A	3	2.39	1.84
	B	1	.60	.46
	AB	3	.55	.42
	Error	192	1.30	
VI. Communication	A	3	1.03	.69
	B	1	.01	.01
	AB	3	1.04	.69
	Error	192	1.50	

TABLE 5.2
(continued)

Criterion		Source of Variance	DF	MS	F
VII.	Reasoning	A	3	3.49	1.56
		B	1	7.22	3.24
		AB	3	1.19	.53
		Error	192	2.23	
VIII.	Process of Thinking	A	3	11.71	3.61*
		B	1	.04	.01
		AB	3	2.93	.90
		Error	192	3.24	
IX.	High Potential	A	3	.45	.32
		B	1	.64	.46
		AB	3	.56	.40
		Error	192	1.40	
X.	Composite Personality	A	3	1.26	1.40
		B	1	.47	.52
		AB	3	1.31	1.46
		Error	192	.90	
XI.	Composite Logic	A	3	1.66	2.27
		B	1	.71	.97
		AB	3	.42	.58
		Error	192	.73	
XII.	Composite Organization	A	3	.24	.67
		B	1	.00	.00
		AB	3	.07	.19
		Error	192	.36	

*F(3,192.05) = 2.65

Subject Fields

Table 5.3 presents the mean criterion ratings assigned to student teachers in the various teaching fields and by type of class. For example, of 257 student teachers observed teaching English a mean rating of 2.58 was given in Criterion IX (High Professional Potential vs. Low Professional Potential). This compares favorably with an overall mean of 2.71 and a mean of 2.67 for all academic-type classes. Only Special Education student teachers were rated higher--2.50.

The high degree of similarity among the means assigned on all criteria regardless of the subject field can be seen. It should also be noted that there is practically no difference in the means assigned to laboratory-type classes and to academic-type classes. By observation it can be seen that the standard deviation is usually at least twice as large as the greatest difference between any means for a given criterion.

Student teachers in Mathematics, followed by those in English, Music, Art, and Special Education ranked highest on Criterion VII (Classroom Discourse Characterized by Reasoning or Creative Thinking). Music, Special Education, and Mathematics teachers were rated highest on Criterion VIII (Directs Attention to Process of Thinking). Foreign Language students received the highest ratings and Social Studies the lowest on Criterion V (Has Mastery of Facts and Organizing Principles of Field).

On Criterion I (Understanding, Friendly), Industrial Education, Science, and Social Studies student teachers were rated lower than others. Special Education teachers were rated the highest. On Criterion II (Planned, Organized, Responsible), student teachers of Business, Foreign Language, and Special Education were rated highest. On Criterion III (Stimulating, Imaginative, Surgent), student teachers of Foreigh Language, Special Education, and English were rated the highest while Industrial Education and Home Economics were ranked the lowest. It should be noted that the difference between laboratory-type classes and academic-type classes was the greatest on this criterion.

On Criterion IV (Perceives Self as Competent), the highest ratings were given in Foreign Language and Mathematics and the lowest in Industrial Education, Home Economics, and Social Studies.

On Criterion VI (Communicates Well and Empathetically), student teachers in Special Education, Foreign Language, and English rank highest with those in Industrial Education, Science, Social Studies and Physical Education lowest.

The differences on criterion ratings were so small among the various subject fields that it is not practical to generalize about a specific subject field.

TABLE 5.3

MEAN ESTIMATES ON EACH CRITERION OF SSTPP TOGETHER WITH STANDARD DEVIATIONS BY SUBJECT FIELDS AND TYPE OF CLASS

	I	II	III	IV	V	VI	VII	VIII	IX
Art N = 65									
M	2.64	2.54	3.10	2.70	2.51	2.93	3.41	4.21	2.62
SD	1.01	1.09	1.07	1.01	.89	1.07	1.58	1.86	.96
Business N = 164									
M	2.50	2.25	3.04	2.62	2.45	2.62	3.71	4.26	2.55
SD	.96	1.05	1.14	1.00	1.11	1.05	1.45	1.74	1.06
English N = 257									
M	2.58	2.36	2.88	2.56	2.55	2.56	3.36	4.19	2.51
SD	.98	1.12	1.21	1.13	1.08	1.12	1.53	1.84	1.07
Foreign Language N = 103									
M	2.54	2.25	2.66	2.23	1.96	2.53	4.01	4.29	2.28
SD	1.01	1.11	1.27	.86	1.02	1.23	1.51	2.00	1.15
Home Economics N = 120									
M	2.62	2.54	3.32	2.84	2.73	2.80	3.87	4.86	2.91
SD	1.00	1.22	1.05	1.17	1.14	1.08	1.52	1.71	1.14
Industrial Education N = 208									
M	2.93	2.72	3.59	2.98	2.77	3.23	3.81	4.40	2.95
SD	.99	1.10	1.14	1.25	1.06	1.10	1.28	1.58	1.10
Mathematics N = 149									
M	2.67	2.42	3.03	2.52	2.32	2.70	3.14	3.73	2.52
SD	1.02	.97	1.11	1.05	1.04	1.06	1.37	1.93	.94
Music N = 137									
M	2.62	2.50	3.25	2.67	2.36	2.80	3.38	3.49	2.62
SD	1.03	.94	1.14	1.11	.89	1.11	1.17	1.18	.97

TABLE 5.3

MEAN ESTIMATES ON EACH CRITERION OF SSTPP TOGETHER WITH STANDARD DEVIATIONS BY SUBJECT FIELDS AND TYPE OF CLASS
(Continued)

	I	II	III	IV	V	VI	VII	VIII	IX
Physical Education N = 289									
M	2.60	2.55	3.11	2.63	2.85	2.88	4.27	4.48	2.77
SD	.98	1.12	1.24	1.13	1.13	1.16	1.61	1.78	1.08
Science N = 140									
M	2.80	2.59	3.23	2.80	2.82	3.06	4.00	4.76	2.90
SD	1.14	1.07	1.29	1.18	1.12	1.24	1.60	1.88	1.11
Social Studies N = 243									
M	2.73	2.68	3.26	2.84	3.06	3.00	3.89	4.83	2.99
SD	1.12	1.07	1.32	1.24	1.25	1.31	1.67	1.83	1.18
Special Education N = 93									
M	2.46	2.28	2.71	2.59	2.51	2.47	3.44	3.65	2.50
SD	1.08	1.07	1.18	1.07	.95	1.15	1.56	1.72	1.04
Laboratory Classes N = 1078									
M	2.64	2.51	3.20	2.72	2.65	2.86	3.78	4.25	2.73
SD	1.01	1.10	1.18	1.13	1.07	1.13	1.48	1.70	1.07
Academic Classes N = 892									
M	2.67	2.48	3.04	2.63	2.63	2.78	3.64	4.39	2.67
SD	1.05	1.08	1.26	1.14	1.18	1.21	1.59	1.91	1.12
Total Group N = 1970									
M	2.65	2.49	3.12	2.68	2.64	2.82	3.70	4.33	2.71
SD	1.03	1.09	1.22	1.14	1.12	1.17	1.54	1.83	1.10

Factor Analysis of SSTPP Ratings

When the criteria were identified it was assumed that there were eight discrete areas of professional behavior and one overarching criterion. Factor analysis using oblique as well as orthogonal matrices revealed only three factors among the nine criteria. The first might be called a personality factor which loaded from Criteria III, VI, I, and IV. The second could be called the logic factor and loaded from Criteria VII and VIII. The third factor could be called organization--loaded from Criteria II and V. The first factor (Personality--subsequently referred to as Criterion X) accounted for 63% of the variance, while the second factor (Logic--referred to as Criterion XI) accounted for 32% of the variance and the third factor (Organization--Criterion XII) accounted for 13% of the variance. Factor scores were computed for each student teacher and are examined in the following sections of this report along with the original nine criteria.

Relationship to Cumulative Grade-Point Average

As shown in Table 5.4, of the 564 student teachers in the study only 510 had cumulative grade-point averages with which to correlate their criterion scores. While correlations for all criteria except VII and VIII are significantly different from zero, they are so small as to be of little practical value for predictive purposes.

Criterion V (Subject Matter) with an r of .26 is the highest. Criterion VI (Communication) is next with an r of .25 followed by the global estimate Criterion IX with an r of .22. Criterion II (Organization), IV (Self-Concept) and III (Surgency) have r's of .17, .16, and .13 respectively. It should be noted that Criteria VII and VIII have no relationship while that for Criterion I (Friendliness) has an r of only .06.

Of the three factor scores, Criterion XII (Composite Organization) correlated the highest with an r of .20, followed by Criterion XI (Composite Logic) with an r of .12 and Criterion X (Composite Personality) with an r of .11.

It is obvious that the criterion estimates are measuring something other than what is measured by academic grades. However, it is interesting that the highest correlations are with Criterion V which rates the student teacher's mastery of his subject field and Criterion VI which rates the student teacher's communicative skills. The criteria concerned with Logic, VII and VIII, pose a serious problem of interpretation. Those students who were rated highest on these criteria were not the "best" academic students--a point which was mentioned in staff meetings as the staff refined its definitions. That Criteria VII and VIII played a relatively minor role in determining the global rating on Criterion IX is shown by the relatively low correlations between Criterion IX and Criteria VII and VIII as reported in Table 5.5.

TABLE 5.4

CORRELATION COEFFICIENTS OF CRITERION RATINGS ON SECONDARY STUDENT TEACHER PERSONALITY PROFILE AND VARIOUS OTHER DATA ON STUDENT TEACHERS

Criterion	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
A. Cumulative Grade-Point Average N = 510	.06	$.17^{**}$	$.13^{**}$	$.16^{**}$	$.26^{**}$	$.25^{**}$	.00	.01	$.22^{**}$	$.11^{*}$	$.12^{*}$	$.20^{**}$
B. Estimated Grade-Point Average N = 299	$.12^{*}$	$.16^{**}$	$.13^{*}$	.04	$.15^{**}$	$.21^{**}$	$.11^{*}$	$.11^{*}$	$.19^{**}$	$.13^{*}$	$.22^{**}$	.10
C. Scholastic Aptitude Test N = 337												
Verbal	.04	-.01	.04	-.07	-.01	.06	-.03	-.00	.08	-.07	.02	.02
Mathematical	-.02	-.01	.01	$-.11^{*}$	-.04	.04	.01	.03	.09	-.04	$.12^{*}$	.05
D. Cornell Critical Thinking Test N = 399	.04	.08	.05	.05	.01	.06	.03	-.02	.02	.07	.08	.00
E. Evaluation by Supervising Teachers N = 423	$.26^{**}$	$.32^{**}$	$.29^{**}$	$.37^{**}$	$.33^{**}$	$.30^{**}$	$.12^{*}$	.05	$.44^{**}$	$.40^{**}$	$.16^{**}$	$.37^{**}$
F. Evaluation by Supervising Teachers Pilot Study N = 55	.11	.33	.32	.31	.23	.05	.26	.35	.37	---	---	---

Signs have been adjusted to eliminate the effects of inverse scoring.

A two-tailed test to determine a correlation coefficient significantly different from zero required $r^{*} \geq .113$ at the .05 level and $r^{**} \geq .148$ at the .01 level when N = 300 pairs.

TABLE 5.5

MATRIX OF CORRELATION COEFFICIENTS-CRITERION IX (HIGH PROFESSIONAL POTENTIAL) AND OTHER CRITERIA

N = 499 Observations

Criterion	Criterion IX
I	.664
II	.731
III	.747
IV	.617
V	.736
VI	.766
VII	.283
VIII	.278
IX	1.000

It is possible that the concept of Thinking which the staff developed and attempted to apply in Criteria VII and VIII is so narrowly conceived as to be unrelated to other qualities or so poorly defined as to be impractical. Or it is possible that the concept of Thinking and the Thinking Process is one that should be identified as a distinct pedagogical skill. In spite of the difficulties encountered in handling this element of teacher behavior the staff is unanimous in its belief that it is important, and is continuing to experiment with observation and reporting techniques.

Relationship to Estimated Grade-Point Average

The estimated grade-point average, computed by the university admissions office for the assistance of counselors, uses scores on the Scholastic Aptitude Test and rank in high school graduating class. Correlations between criterion ratings and estimated grade-point average, like those with actual cumulative grade-point average, are statistically significant but are so low as to have little predictive value. As shown in Table 5.4, interestingly, Criterion XI (Composite Logic) has the highest correlation, r = .22, while Criterion IV (Self-Concept) has a low of r = .04. Except for Logic, the correlations are similar to those found for cumulative grade-point average.

Relationship to Scores on Scholastic Aptitude Test

The fact that criterion assessments fail to show any relationship whatsoever with scores on the SAT tests must be viewed as significant. If one assumes that the SAT test gives a reasonable estimate of a student's verbal and mathematical abilities it is apparent that these same abilities did not play an important role in influencing the observer's judgment. The SSTPP criterion scores reflect something other than these abilities. The slight positive correlations on the Global Criterion IX suggest that generalized ability as perceived by the observer plays a more important part in making the global estimate than in making estimates on the other discrete criteria.

Criterion XI (Composite Logic) which correlates with SAT Mathematics score at .12 and the fact that Mathematics teachers obtained the highest score on Criterion VII (Reasoning) as reported in Table 5.3 suggest some relationship between mathematical ability and the use of Logic in the classroom.

Relationship to the Cornell Critical Thinking Test

The Cornell Critical Thinking Test was selected because its author had been associated with the University of Illinois group under the leadership of B. Othanel Smith concerned with studying logic and teaching.[3] A distribution of raw scores is included as Figure 5.3.

The correlations between scores on the Cornell Critical Thinking Test and criterion estimates as presented in Table 5.4 are so low as to be of no significance. It had been hypothesized that student teachers who were rated high on Criterion VII (Classroom Discourse Characterized by Reasoning or Creative Thinking) and Criterion VIII (Directs Attention to Process of Thinking) and on Criterion XI (Composite Logic) would score high on the Critical Thinking Test. The hypothesis had to be rejected because none of the correlation coefficients were significantly different from zero. Apparently SSTPP scores even on criteria related to logic reflect some other quality than that measured by the Cornell test.

It is interesting to note that almost the same absence of relationship exists with respect to SAT scores and criterion estimates.

The scores on the Cornell Critical Thinking Test positively correlated with cumulative grade-point average, r = .35 (N = 290) and scores on the Scholastic Aptitude Test, r = .18 for the Verbal Test and r = .19 for the Mathematics Test (N = 265).*

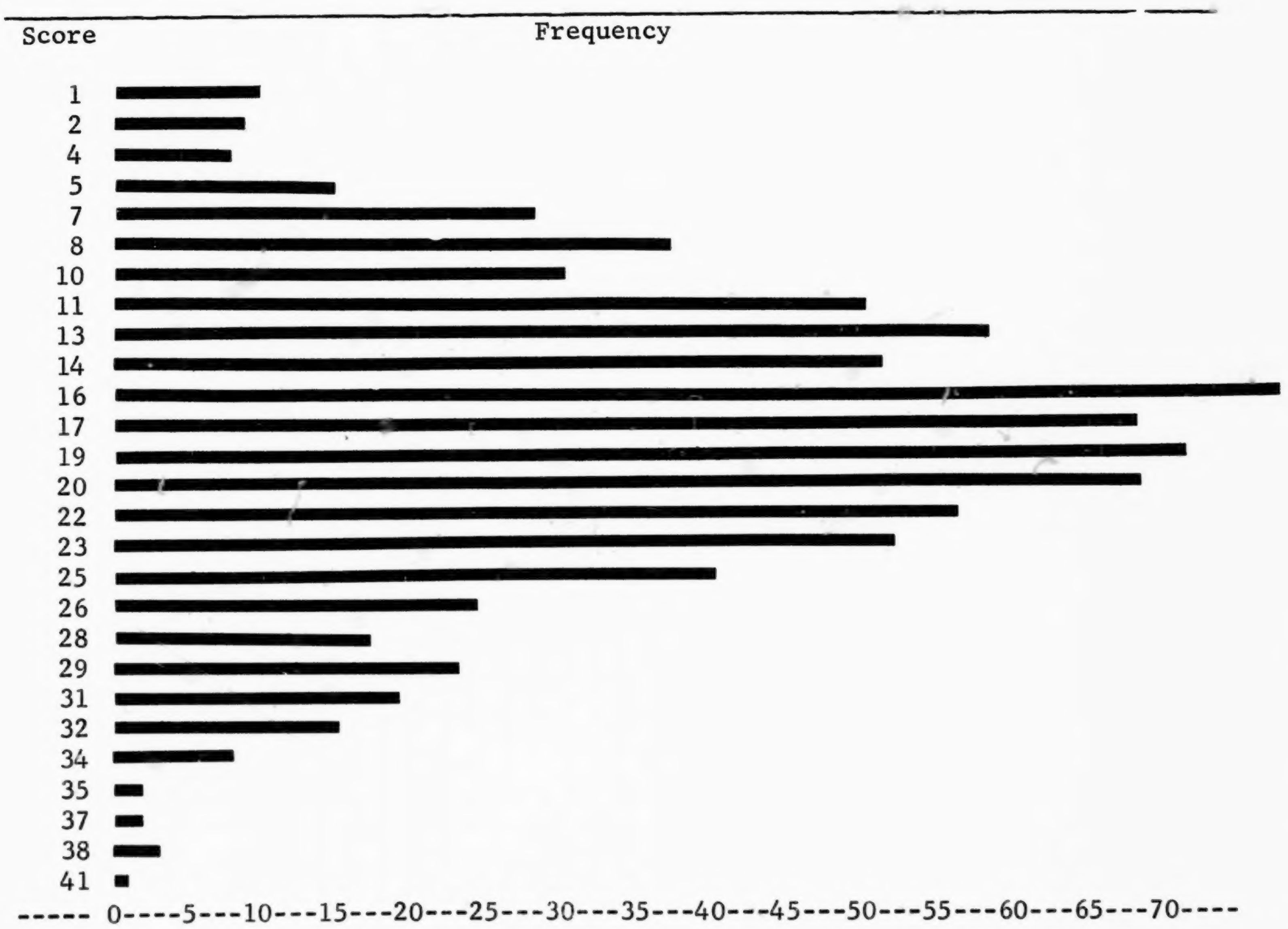
---

[3] Robert H. Ennis, The Development of a Critical Thinking Test, Ph.D. Dissertation, University of Illinois, 1958 (University Microfilms, Inc., Ann Arbor, 1962), p. 156.

*See the Master Matrix of Correlation Coefficients between Assessments on the Twelve Criteria and Other Variables. Appendix D.

FIGURE 5.3

HISTOGRAM OF RAW SCORES ON CORNELL CRITICAL THINKING TEST



Mean = 17.69 Standard Deviation = 7.32 Number = 836
Skewness, which describes the degree of asymmetry of a distribution, = 6.128.
Kurtosis, which describes the relative flatness or peakedness of a distribution = 10.816.
Score = Rights minus one-half wrongs.

Only two significant correlations appear between scores on the CCTT and the 16 PF test. A positive r .24 was obtained on Factor B--General Intelligence, Bright and a negative r -.13 was obtained on Factor H--Adventurous vs. Shy, Timid. Both of these correlations were significantly different from zero at the .01 level. This adds a bit more evidence that student teachers who score high on Criteria VII and VIII fit the stereotype of the quiet, timid intellectual.

It seems strange that no significant relationships were found between scores on the CCTT and the time distribution reported in TCAP. Although the highest correlation (r = .09) was found between CCTT and the amount of time devoted to the Thinking Process and the next highest (r = .07) between CCTT and time devoted to Logical Thinking, these relationships failed to be of statistical significance. Similarly, scores on the CCTT were not related to evaluations made by cooperating teachers.

Further study of the CCTT is necessary before its value in a teacher education program can be determined.

Relationship to Evaluation by Supervising Teachers

Student teachers at Indiana State University are not assigned letter grades. The official transcript reports credit or no credit. Each supervising teacher or cooperating teacher does complete an evaluation form which includes a check list of twenty-two items plus three open-ended statements, and a summary evaluation as follows: "In my judgment this student will become a teacher who is (check one only) Outstanding, Above Average, Average, or Below Average."

Correlations between ratings so given by cooperating teacher and criterion scores assigned by Division of Teaching staff members are shown in Table 5.4.

It should be noted that these are the highest correlations obtained in any comparisons. With the exception of the criteria concerned with logic, the correlations are not only significantly different from zero but have some moderated predictive value in at least two instances.

A correlation coefficient of .44 between Criterion IX (Professional Potential) and the supervisors' estimate of the students' future success bespeaks some agreement among evaluators. Criterion X (Composite Personality) which correlates at the .40 level with supervising teachers' judgment suggests that personality does play a most important role in teacher evaluation. This point is also supported by the r of .37 which shows the relationship between scores on Criterion IV (Perceives Self as Competent) and cooperating teachers' judgments.

Criterion XII (Composite Organization) plays an important role in supervising teachers' judgments as shown by r of .37 for Criterion XII which is made up of Criterion II (Organized, Planned) with an r of .32 and Criterion V (Mastery of Subject Matter) with an r of .33.

Again, the criteria concerned with Thinking and Logic fail to show any relationships.

The fact that criterion scores do correlate at a moderate level with estimates made by cooperating teachers whereas they fail to correlate with other descriptive data suggests that the criterion scores are reflecting some specific skills or competencies which could be identified as pedagogical or teaching behaviors and are similar to the qualities identified by cooperating teachers.

In a pilot study conducted in 1964, as reported in Table 5.4, cooperating teachers' evaluations were correlated with the ratings on SSTPP. These coefficients are very similar to those obtained in the study proper on Criteria II, III, IV, V, and IX. In the pilot study, Criterion VI (Communication) had a low correlation, r = .05, while Criteria VII and VIII (Logic) had relatively high correlations, r = .26 and .35. No explanation of the low correlation on communication is apparent. The surprisingly high correlations on the logic criteria are probably due to the fact that the staff was involved in designing the instrument and writing the Guide at that time. In the process of identifying student teacher behaviors the students and their cooperating teachers were consulted. There was likely an unconscious tendency to "teach" for high performance on Criteria VII and VIII. During the two years of the project supervisors avoided any action which would encourage students to strive for a special rating on SSTPP. While the TCAP observation was discussed with students, the ratings on SSTPP were not.

Similar results were obtained in a small project described as Treatment 5 in Chapter III. During a summer session when supervisors actually taught Education courses to the students they supervised, improvement in both student performance and observer evaluation occurred on Criteria VII and VIII.

While it had been hoped that the careful definition of the criteria in terms of teaching behavior and controlled methods of observation would result in estimates which would be related to other data commonly used to predict teacher performance, the results failed to reveal any such relationship.

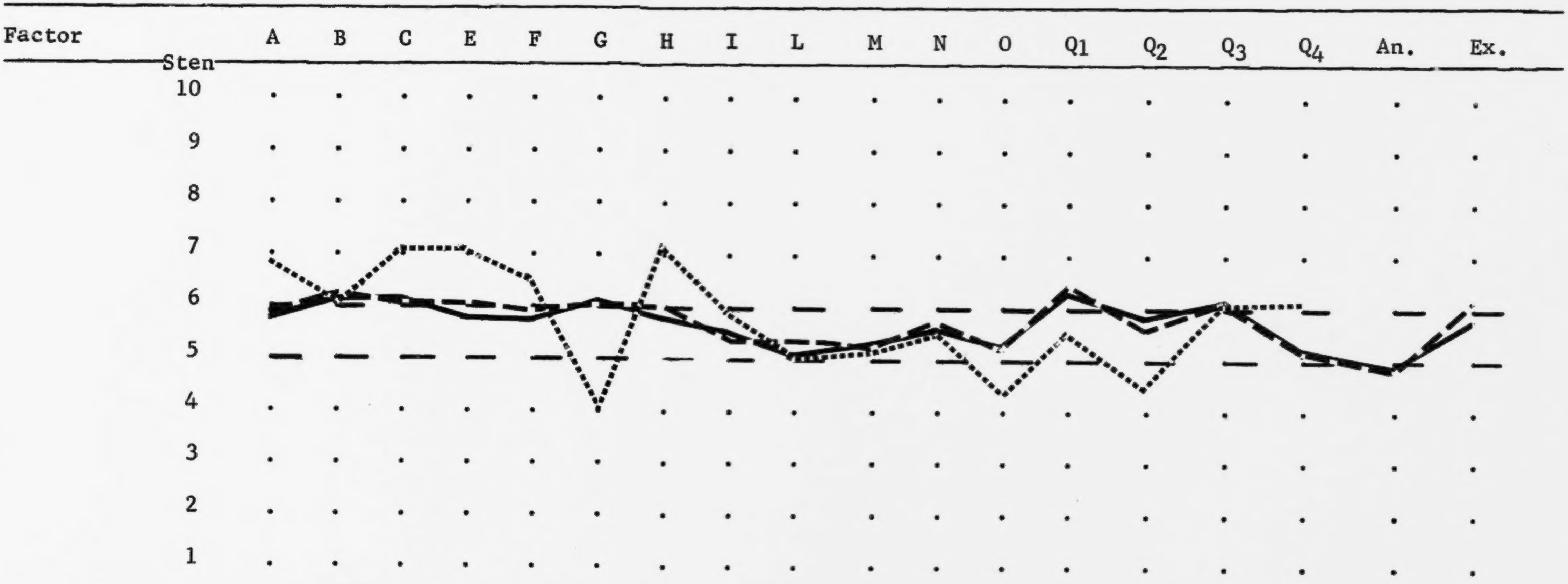
Relationship to Sixteen Personality Factor Questionnaire

As shown in Figure 5.4, the two profiles made by the mean scores on the 16 PF Questionnaire are almost identical between the Experimental Group and student teachers not included in the Experimental Group. Both groups are slightly above the mean on Factors B, C, G, Q_1 and Q_3 and lower on the second order factor of Anxiety. With respect to the other factors the means fall within the interval between Sten 5 and Sten 6 which by definition includes scores one-half of a standard deviation below and above the mean.

The profile differs from that provided by R. B. Cattell for 1128 college undergraduates. These undergraduates were considerably higher on Factors A, C, D, and H and lower on Factors G, Q_1, and Q_2 than either of the student teacher groups at Indiana State University.

FIGURE 5.4

DISTRIBUTION OF MEANS AND STANDARD DEVIATIONS OF STEN SCORES ON 16 P.F.
For Experimental Group, Non-Experimental Group, and College Undergraduates



		A	B	C	E	F	G	H	I	L	M	N	O	Q_1	Q_2	Q_3	Q_4	An.	Ex.
Experimental Group	Mean	5.78	6.09	6.12	5.74	5.72	6.05	5.78	5.55	5.12	5.39	5.62	5.28	6.31	5.78	6.06	5.20	4.97	5.82
N = 425 ▬▬	S.D.	1.76	1.96	1.75	1.94	1.96	1.96	1.91	1.91	1.98	1.86	2.08	1.88	1.89	1.85	1.83	1.75	1.83	2.10
Non-Experimental Group	Mean	5.84	6.26	6.02	6.02	5.85	6.07	6.00	5.59	5.07	5.33	5.75	5.28	6.42	5.53	6.09	5.05	4.86	6.14
N = 192 ━ ━	S.D.	1.85	1.84	1.82	1.88	2.05	1.84	2.07	1.92	1.82	1.81	1.85	1.95	1.91	2.03	1.86	1.89	1.89	2.31
College Undergrads N = 1128 ┅┅	Mean GPSS *	6.8	6.1	7.3	7.2	6.5	4.1	7.6	5.9	5.2	5.7	5.5	4.4	5.7	4.5	6.1	6.2	--	--

* Reported in R. B. Cattell, Handbook for the Sixteen Personality Factor Questionnaire, 1957 Edition with 1964 Supplementation, p.28.

TABLE 5.6

CORRELATION COEFFICIENTS OF CRITERIA RATINGS
AND SCORES ON THE SIXTEEN PERSONALITY FACTOR QUESTIONNAIRE

Criteria	Warm	Bright	Ego Strength	Dominance	Surgency	Conscientious	Adventurous	Sensitive	Suspecting	Unconventional	Sophisticated	Timid	Radicalism	Self-Suffecient	Controlled	Tense	Anxiety	Extroversion	Multiple--r.	F Value
	A	B	C	E	F	G	H	I	L	M	N	O	Q1	Q2	Q3	Q4				
I. Warm, Friendly	.02	.00	.10*	.01	.04	.09*	.09*	.02	-.04	.01	-.07	-.01	.02	-.03	.11	-.06	-.10*	.07	.22	1.13
II. Planned, Organized Responsible	-.03	.02	.16**	.01	.00	.16**	.08*	-.05	-.08	-.05	.04	-.11**	.06	.02	.11**	-.14**	-.17**	.03	.23	1.20
III. Stimulating, Surgent	-.04	-.01	.16**	.17**	.07	.08*	.24**	.02	-.03	.05	-.01	-.06	.04	-.03	.07	-.08	-.14**	.18**	.31	2.46**
IV. Self-Concept	-.03	-.01	.14**	.17**	.10**	.16**	.21**	.10*	-.01	-.02	.00	.15**	.04	-.04	.08	-.07	-.17**	.19**	.33	2.78**
V. Subject Mastery	-.04	.03	.13**	.06	.04	.16**	.13**	-.01	-.05	-.02	-.04	-.11**	.08	-.04	.10*	-.12**	-.15**	.09*	.26	1.64*
VI. Communication	.03	.05	.17**	.17**	.07	.11*	.19**	.02	-.04	-.02	-.03	-.10*	.04	-.04	.11*	-.10*	-.16**	.18**	.30	2.20**
VII. Reasoning--Logic	-.13**	-.02	.00	-.03	-.11*	.05	-.08	.03	-.06	.04	-.03	-.02	.08	.08	.01	-.04	-.02	-.11*	.22	1.12
VIII. Process of Thinking	-.06	-.07	-.01	-.04	-.14**	.14**	-.08	.06	-.07	.05	-.09	-.02	.02	.03	.06	-.06	-.03	-.11*	.29	2.00**
IX. Professional Potential	.05	.04	.15**	.09	.05	.13**	.15**	-.04	-.05	-.01	-.01	-.11	.07	-.02	.12**	-.14**	-.17**	.11*	.26	1.66*
X. Composite Personality	.01	.02	.14	.14	.09	.09	.23	.00	.06	.01	.01	.05	.06	.08	.15**	.10	-.17**	.19**	.30	1.48
XI. Composite Logic	-.17*	-.02	.09	.06	-.05	.10	-.05	.05	-.01	.07	-.10*	-.01	.06	.11*	.01	-.05	-.03	-.06	.32	1.71*
XII. Composite Organization	-.06	.03	.05	-.03	-.04	.13**	-.02	-.08	-.05	-.10	.02	-.08	.09	.00	.13**	-.08	-.08	-.04	.24	0.94

*r≥.05 level
**r≥.01 level

A two-tailed test to determine a correlation coefficient significantly different from zero required r≥.088 at the .05 level and r≥.115 at the .01 level.

One of the contributing factors to the low correlations between SSTPP criterion ratings and scores on the 16 PF test is the tendency of scores on the 16 PF to cluster close to the mean.

The similarity of several terms used in the SSTPP and the 16 PF Questionnaire suggested possible relationships. For example, Criterion I was defined as Understanding, Friendly vs. Egocentric, Aloof which would appear to compare with Factor A which is described as Cyclothymia (Warm, Sociable) vs. Schizothymia (Aloof, Stiff).

Criterion II, which was defined as Planned, Organized, Responsible vs. Unplanned, Disorganized, Irresponsible, would seem to compare with Factor G--Conscientious, Persistent vs. Casual, Undependable. Criterion III (Stimulating, Imaginative, Surgent vs. Dull, Routine, Unimaginative) would seem to compare with Factor F--Surgency--Enthusiastic, Happy-Go-Lucky vs. Desurgency--Glum, Sober, Serious, and Factor H--Adventurous--Active, Responsive, Genial vs. Shy, Timid. Criterion IV (Perceives Self as Competent vs. Perceives Self as Less than Adequate) would appear to have some things in common with Factor C--Emotional Stability or Ego Strength and Factor H--Adventurous vs. Shy, Timid and Q_2--Self-Sufficiency.

The factor analyses made of the nine criteria identified one major factor which was called Criterion X (Composite Personality) and two minor factors, Criterion XI (Composite Logic) and Criterion XII (Composite Organization). One might expect the criteria which loaded the Composite Personality Factor, namely I, III, IV, VI and IX, to be related to scores on the 16 Personality Factor Test.

It was hypothesized that high criterion ratings would correlate with similarly high scores on the comparable 16 PF Questionnaires.

As a matter of fact, many correlations were found to be significantly greater than zero but none of them were high enough to be useful as predictors. They do suggest a possible relationship and argue for further study. It should be kept in mind that the criteria were conceived of as performance criteria rather than personality criteria. Observers made their estimates on the basis of what the teacher did rather than what the teacher "was." The criteria reflect a theory of teaching while the 16 PF scores reflect a theory of personality.

Criterion I (Understanding, Friendly) as shown in Table 5.6 correlated not with Factor A but slightly significantly at the .05 level with Factor C, r = .10; Factor G, r = .09; Factor H, r = .09; and with the second order factor of Anxiety, r = -.10.* An examination of the glossary of Operational Definitions for SSTPP and the illustrative anecdotal reports would explain the relationship of Criterion I and Factors C, G, H, and Anxiety, as follows: the student teacher who scores high on Criterion I is Attentive to People (G), Friendly rather than Shy, Withdrawn, or Aloof (H), and low on the second-order factor, Anxiety. The concept of Emotional Stability, Mature and Calm, which described Factor C, does not appear in any definitions of Criterion I.

---

* A two-tailed test of correlation coefficent being significantly different from zero requires $r > .088$ at the .05 level and $r > .115$ at the .01 level.

Criterion II (Planned, Organized, Responsible) did correlate, r = .16, with Factor G--Conscientious which is significant at the .01 level. Similar relationships are suggested with Factors C--Emotional Stability, r = .16; O--Confidence, r = .11; Q_3--Controlled, r = .11; and Q_4--Tense, r = .14; and the second-order factor of Anxiety where a negative relationship, r = -.17, was found. A correlation significant at the .05 level was found with Factor H--Adventurous, r = .08.

Descriptive data for Criterion II emphasized the planning and organizing ability of the student which does include the conscientiousness of Factor G. It is interesting that other factors are related to Criterion II scores in the same degree: Factor C--Emotional Maturity; the second-order factor--Anxiety; Factor O--Self-Confidence as contrasted with Timidity; Factor Q3--Controlled; and Q_4--Tense. These may contribute to the judgments given but they were not anticipated.

Criterion III (Stimulating, Imaginative, Surgent vs. Dull, Routine, Unimaginative) failed to relate significantly to Factor F--Surgency by a small amount, r = .082, while an r of .088 was required for significance at .05 level. However, it was related to Factors H--Adventurous, r = .24; E--Dominance, r = .17; and C--Ego Strength, r = .16. Student teachers who were rated high on Criterion III (Stimulating, Imaginative, Surgent) received significant scores on both second-order factors. For Anxiety the relationship was negative, r =-.14; for Extroversion it was positive, r = .18.

On Criterion IV (Perceives Self as Competent vs. Perceives Self as Less than Adequate) the observers were rating a generalized personality trait rather than a pedagogical skill. The ratings were significantly related at the .01 level on nine of the eighteen factors. Student teachers who were rated high on self-concept scored high on Factors H--Adventurous, r = .21; E--Dominance, r = .17; G--Conscientious, r = .16; O--Timid, r = -.15; C—Ego Strength, r = .14; I--Sensitivity, r = .10; F--Surgency, r = .10; Anxiety, r = .17; and Extroversion, r = .19.

Criterion V (Has Mastery of Facts and Organizing Principles of Field vs. Has Only Minimum Knowledge of Field) was not considered a personality criterion. However, it was related to six of the factor scores at the .01 level and two at the .05 level. Student teachers who were rated high in subject mastery tend to rate low in Anxiety, r =-.15 and Factor O--Timid, r = -.11, and high in Factor G--Conscientious, r = .16; C--Ego Strength, r = .13; H--Adventurous, r = .13; and Q_4--High Tension, r = .12. At the .05 level they were Q_3--Controlled, r = .10 and Extroverted, r = .09.

Criterion VI (Communicates Well and Empathetically vs. Communicates Ineffectively and Perfunctorily) was related at the .01 level to Factor H--Adventurous, r = .19; Extroversion, r = .18; E--Dominance, r = .17; C--Ego Strength, r = .17; and Anxiety, r =-.16. At the .05 level communication skill seemed to be related to Q_3--Controlled, r = .11; G--Conscientious, r = .11; Q_4--High Tension, and O--Timid, r = -.10.

Criterion VII (Classroom Discourse Characterized by Reasoning or Creative Thinking vs. Classroom Discourse Characterized by Simple Recall) is related negatively to three personality factors: A--Warm, Sociable, r = -.13; F--Surgency, r = -.11; and Extroversion-Introversion, r = -.11. Apparently the concept that students who rate high on reason and logic have to be cold, unexciting, and introverted persists in the rating of student teachers.

Criterion VIII (Directs Attention to the Process of Thinking vs. Fails to Call Attention to the Process of Thinking) is similarly negatively related to Factor F--Surgence, r = -.14 and Extroversion, r = -.11. However, it is positively related to G--Conscientious, r = .14.

Criterion IX (High Professional Potential vs. Low Professional Potential) is related to seven factors: Anxiety, r = -.17; H--Adventurous, r = .15; C--Ego Strength, r = .15; Q_4--High Tension, r = -.14; G--Conscientious, r = .13; Q_3--Controlled, r = .12; and Extroversion, r = .11.

Criterion X (Composite Personality) is related to Factors C--Ego Strength, r = .14; E--Dominance, r = .14; Q_3--Controlled, r = .15; Extroversion, r = .19; and Anxiety, r = -.17.

Criterion XI (Composite Logic) is negatively related to Factor A--Warm, Sociable, r = -.17 and Factor N--Sophisticated, r = -.10, and positively related to Q_2--Self-Sufficient, r = .11.

Criterion XII (Composite Organization) is related to Factors G--Conscientious, r = .13 and Q_3--Controlled, r = .13.

Because the 18 separate factors had only a low correlation to the criterion scores, a multiple regression analysis was run on the 16 PF scores. Table 5.6 includes in the right-hand column the values of the multiple r plus the corresponding F values. The multiple r was related at the .01 level to Criteria III, IV, VI, and VIII. It was related at the .05 level to Criteria V, IX, and XI.

In summary, the relationships between criteria estimates and scores on the 16 PF test are statistically significant. It must be pointed out that the correlations are of such a low value that their usefulness as predictors is not very great. However, it must also be recognized that the relationships are all consistent with the theoretical constructs of the two instruments. While the evidence may not be conclusive, there is certainly a strong suggestion that the criteria are not inconsistent with some scores on the 16 PF Questionnaire.

Relationships Between Assessments Made on an Initial Interview and Those Made from Classroom Observations

Prior to assigning a student teacher to a cooperating school, personal interviews were conducted by staff members and Departmental Supervisors. The actual assignment was decided upon in a joint conference of staff and Departmental members. For purposes of this study the staff members made an estimate of the student's future performance on the nine criteria. For one semester Departmental Supervisors estimated the student's performance on the global Criterion IX.

Approximately 30 minutes were devoted to the interview in which the staff member reviewed the student's written application, his college record and a four-page personal data sheet reporting the student's background, his non-school experiences, his interests and his aspirations. Considerable time was devoted to exploring the type of school and teacher to which the student should be assigned. While some attention was given to university policies regarding assignments, the interview was structured to enable the staff member to become acquainted with the student.

At the end of the interview, the student's predicted performance on each of the nine criteria was recorded using the seven-point scale. All assessments were made without reference to any previous ratings.

Five hypotheses were posed with respect to the interview:

> First, ratings based on the interview would correlate with subsequent ratings based on observed classroom behavior.
>
> Second, the interview itself would provide clues from which to predict the student's ratings on Criteria I (Friendliness), III (Surgence), and IV (Self-Concept) and the predictions would yield the highest correlations.
>
> Third, clues could be obtained from the student's records and the information he provided which would result in reliable predictions for Criteria II (Organization) and V (Subject Matter).
>
> Fourth, the lowest correlations would be found on Criteria VII (Logic) and VIII (the Thinking Process).
>
> Fifth, the correlations would be higher when the interview and the observation were performed by the same person.

Table 5.7 reports the correlations among the criteria ratings made in the initial interview and those made in two subsequent observations of the student's performance in class. Columns 1 and 2 show the correlations for those instances when the same staff member conducted the interview and made the observations. Columns 3 and 4 report the correlations for those instances when one person conducted the interview and another staff member made the observations. Column 5 shows the corre-

TABLE 5.7

CORRELATION COEFFICIENTS OF CRITERION RATINGS MADE FROM INTERVIEWS AND OBSERVATIONS

Criterion	By Same Person N = 129		By Different Persons N = 511		Dept. Supervisor Global Estimate N = 52
	First Obs.	Second Obs.	First Obs.	Second Obs.	Second Obs.
I	.23	.18	.16	.14	.08
II	.36	.28	.17	.15	.41
III	.35	.35	.26	.22	.30
IV	.27	.28	.20	.22	.40
V	.33	.24	.22	.22	.40
VI	.28	.40	.13	.20	.31
VII	.00	.06	.11	.10	.03
VIII	.03	.07	.26	.05	.08
IX	.34	.34	.21	.19	.37

lations between the single rating given by the Departmental Supervisor and ratings on each criteria given by the Division of Teaching staff person on his second observation.

While the correlations are all relatively low they are significantly greater than zero. It is obvious that the fifth hypothesis was confirmed in the main: when the same person conducts the interview and makes the observation his two estimates correlate more closely than when one person conducts the interview and another makes the observation. The only exception involves the criteria related to logic and then the differences are so slight as to be insignificant. No explanation for the relatively high correlation (.26) on Criterion VIII between the interview and the first observation when made by a different person is apparent. The correlation with the second observation drops back to .05.

One can not completely reject the first hypothesis, although the evidence is not conclusive. Even the highest correlation of .40 has an index of forecasting efficiency of only 8.4 percent. While the correlations are relatively higher on the personality criteria and the subject-matter-organization criteria than on those related to logic, they are not high enough to warrant placing much confidence in an interview as a predictor of teaching behavior.

There is some evidence for accepting the third hypothesis that information on subject matter and organization can be obtained in an interview. Most of the correlation on Criteria V (Subject Matter) and II (Organization) are relatively high. This is especially true for Departmental Supervisors.

The second hypothesis that the interview can predict subsequent classroom behavior must be rejected. When one eliminates the logic criterion, the lowest correlation was obtained for Criterion I (Friendliness). Apparently this aspect of personality appears differently in an office interview and a classroom teaching situation. Criterion III (Surgency) and Criterion IX (Professional Potential--global rating) have the highest composite correlations.

For one semester during the project Departmental Supervisors, who interview students as part of the assignment procedure, estimated the student's success as a student teacher on a seven-point scale. In effect this was a rating on Criterion IX. Column 5 of Table 5.7 shows the correlations between these ratings and criterion scores given by the staff for the second observation. While these correlations are too low to be of practical value they are in the main somewhat higher than the correlations obtained when the staff estimated each specific criterion. Again correlations on the logic criterion were negligible as was the correlation for Criterion I (Friendliness). Apparently Departmental Supervisors perceive the following criteria as important: II (Organization), IV (Self-Concept), V (Subject Matter) and, to a lesser degree, VI (Communication).

The correlation between Departmental Supervisor estimate and the student's grade-point average was .57 as compared with a .37 for that of the staff. This suggests that Departmental Supervisors were more influenced by the student's academic record than were Division of Teaching staff.

That interviews are not very reliable predictors of classroom behavior should not be surprising in the light of research conducted largely in industry. Hollingsworth's classical study found an almost complete lack of agreement when twelve sales managers ranked fifty-seven applicants on the basis of an interview.[4]

The low correlation found between criterion estimates and other data descriptive of the students is consistent with other studies.

Marsh and Wilder in summarizing some 900 reports of research on effective instruction reported that, among others, none of the following had value as a predictor of rated instructor competence: intelligence, education, scholarship indicated by grade-point average, knowledge of subject matter, professional information, and personality tests.[5]

---

[4] H. L. Hollingsworth, Judging Human Character (New York: Appleton-Century, 1922).

[5] Joseph E. Marsh and Eleanor W. Wilder, Identifying the Effective Instructor: A Review of the Quantitative Studies, 1900-1952 (San Antonio, Texas: Air Force Personnel and Training Research Center, Lackland Air Force Base, 1958), p. 6.

CHAPTER VI

CONCLUSIONS AND IMPLICATIONS

Conclusions

The nine criteria provided an adequate rubric for evaluating student teacher performance. Professionally qualified University supervisors were able to make evaluative judgments of student teacher behavior on the nine evaluative criteria by observing the student teacher in the classroom with sufficient reliability for the rating to be meaningful and useful. The reliability of the appraisals was improved by special training.

Less difficulty in making evaluative judgments was encountered with respect to Criteria I, II, III, IV, V, VI and IX than with Criteria VII and VIII.

Student teachers involved in this study rated the lowest on Criteria VII and VIII which are concerned with reasoning, logic and the process of thinking, and highest on Criteria II, IV, V and VI which are concerned with organization and personality.

Student teachers at Indiana State University devoted only a minor fraction of their classroom activity to higher order intellectual processes.

The correlations between estimates on the nine criteria and other data were so low as to be of little predictive value. The highest correlation, which was with Cumulative Grade-Point Average, had an index of forecasting efficiency of only four percent. No relationships were found with scores on the Scholastic Aptitude Test, the Cornell Test of Critical Thinking, or TCAP records of the way the student spent his time in the classroom. These findings are consistent with those of past research which has been described in these terms: ". . . the majority of educational researchers have been saying for the last decade: we do not know how to define, prepare for, or measure teacher competence."[1]

Sufficient relationships with slight predictive value were found with some scores on the 16 PF Questionnaire to suggest the value of further study.

Variations in *observational practices such as timing, length or* frequency do not materially influence reliability.

Interviews are not satisfactory predictors of student teacher classroom behavior. Global estimates of professional potential made by trained college supervisors are significantly and positively related to similar predictions made by the cooperating supervising teachers in the public schools.

While the nine criteria were assumed to consist of eight discrete elements and one composite element, factor analysis revealed only one major factor identifiable as "personality," a minor factor identifiable as "logic" or "thinking," and a minor factor identifiable as "organization."

---

[1] Bruce J. Biddle, op. cit. p.3

Implications

When this project was inaugurated, the investigators believed that one of the chief purposes of education was the improvement of the student's ability to think. Three years of investigation have reinforced this belief and revealed some of the difficulties involved in dealing with a concept as complex as thinking. The study has documented the small amount of classroom time devoted to thinking, and suggests that teachers need to be not only committed to the scientific method but also equipped with those pedagogical skills which will enable them to help others increase the areas of life subject to rational control. This is not to deny the important role played by aesthetics, emotions and the non-rational.

The definition of Thinking agreed upon attempted to be inclusive of all higher levels of intellectual activity without becoming doctrinaire. As explained in Chapter II, the study was concerned with discovering ways in which the teacher's role in the Thinking Process could be identified regardless of the teacher's background; that is, whether the teacher may have been influenced by formal logic, the five steps in the scientific method, the ideas popularized in propaganda analysis, the concepts associated with General Semantics, the Socratic Method, the modern method of inquiry, the "New Think," the research approach, or simple chance encounters with ways of discovering truth.

The student teacher's need, identified in this study, for the intellectual skills was echoed in an observation by A. M. Tibbetts, ". . .many students are failing to learn the art of consecutive, logical thinking and writing. On most issues of importance that arise in university life, students are failing to investigate fully, clarify premises, define terms, think logically, use evidence properly and write (or speak) precisely, truthfully and to the point."[2] He goes on to propose a new college course called "Investigating, Thinking, and Arguing."

In the past, key words in teacher education included life adjustment, human relations, democracy, core, general education and individualized instruction. The current educational key words include relevance, role differentiation, interaction, media and programmed instruction. This study suggests the need to emphasize the pedagogical processes related to thinking. The study has been of practical use in alerting student teachers to the potentials of improving the intellectual level of classroom discourse. It has indicated the need for still more effective instructional material and better teaching techniques for helping students think.

While this study has made some progress in developing concepts, instruments and techniques for recording and rating student teacher classroom behavior, further development and refinement are needed.

---

[2] A. M. Tibbetts, "To Encourage Reason on the Campus: A Proposal for a New College Course in Thinking and Writing," American Association of University Professors Bulletin 54.4 (Dec. 1968), pp.466 ff.

In accepting the paradigm $P_t \rightarrow D_t \rightarrow R_t \rightarrow P_p \rightarrow D_p \rightarrow R_p \rightarrow P_t \rightarrow$ etc., where P = perceiving, D = diagnosing and R = reacting, subscripts t = teacher and p = pupil,[3] as an appropriate analysis of teaching, one recognizes that interaction is a necessary, but not sufficient, condition for teaching success. There must be interaction but it alone is not enough. The keenness of the perception, the accuracy of the diagnosis and the range of the reactions will determine quality. The criteria used in this study are useful in evaluating the quality of the teacher's classroom behavior.

---

[3] B. O. Smith, "Concept of Teaching," *op*. *cit*.

BIBLIOGRAPHY

Amidon, E. J., and Ned A. Flanders. *The Role of the Teacher in the Classroom*. Minneapolis: Paul S. Amidon & Associates, 1963.

Aschner, Mary Jane. *The Analysis of Classroom Discourse--A Method and Its Uses*. Unpublished doctoral dissertation. Urbana, Ill.: University of Illinois, 1958.

Aschner, Mary Jane. "The Language of Teaching," *Teachers College Record*. Vol. 61, February, 1960.

Aschner, Mary Jane and others. *A System for Classifying Thought Processes in the Content of Classroom Verbal Interaction*. Cooperative Research Project No. 965 of the U. S. Office of Education. Urbana, Ill.: Institute for Research on Exceptional Children, University of Illinois, 1965.

Bellack, Arno, and Joel R. Davitz. *The Language of the Classroom*. New York: Institute of Psychological Research, Teachers College, Columbia University, 1963.

Bellack, Arno A., Ronald T. Hyman, Frank L. Smith, Jr. and Herbert Kliebard. *The Language of the Classroom*. Meanings Communicated in High School Teaching, Part II. Cooperative Research Project No. 2023, 1965.

Benne, Kenneth D., George E. Axtelle, B. Othanel Smith and R. Bruce Raup. *The Discipline of Practical Judgment in a Democratic Society*. 1943. Reprinted as *The Improvement of Practical Intelligence*. New York: Harper and Bros., 1950.

Berman, Louise M. (ed.) *The Nature of Teaching*. Milwaukee, Wisc.: School of Education, University of Wisconsin at Milwaukee, 1962.

Berry, John R. *Professional Preparation and Effectiveness of Beginning Teachers*. Coral Gables, Fla.: Graphic Arts Press, 1960.

Biddle, Bruce J., and William S. Ellena. *Contemporary Research on Teacher Effectiveness*. New York: Holt, Rinehart and Winston, 1964.

Bills, Robert E. *About People and Teaching*. Bulletin of the Bureau of School Service, Vol. XXVIII, No. 2. Lexington: University of Kentucky, December, 1955.

Bloom, Benjamin, (ed.) *Taxonomy of Educational Objectives*. New York: Longmans, Green, 1958.

Broudy, Harry S., B. Othanel Smith, and Joe R. Burnett. *Democracy and Excellence in American Secondary Education*. Chicago: Rand McNally, 1964.

Brown, Bob Burton. "Observer-Judge Ratings of Teacher Competence," *Childhood Education*. Vol. 44, November, 1967.

Brown, Bob Burton. *The Experimental Mind in Education*. New York: Brown, Harper & Row, 1968.

Bruner, Jerome. *The Process of Education*. Cambridge, Mass.: Harvard University Press, 1962.

Bush, Robert N., Norman J. Boyan, and Dwight W. Allen. *Stanford University School of Education Secondary Teacher Education Program*. 1965-1966. Resident Supervisors Handbook. Palo Alto. (Offset.)

Cohen, Morris R., and Ernest Nagel. *An Introduction to Logic and the Scientific Method*. New York: Harcourt, Brace, 1934.

Combs, Arthur W. (ed.) *Perceiving, Behaving, Becoming*. Yearbook, 1962. Washington, D. C.: Association for Supervision and Curriculum Development, 1962.

Combs, Arthur W. *The Professional Education of Teachers*. Boston: Allyn & Bacon, 1965.

Conant, J. B. *The Education of American Teachers*. New York: McGraw-Hill, 1963.

Cyphert, Fredrick R., and Ernest Spaights. *An Analysis and Projection of Research in Teacher Education*. USOE Cooperative Research Project No. F-015. Columbus, Ohio: The Ohio University Research Foundation, 1964.

Fall, Charles R., *et al*. *Individualized Instruction in Teacher Education*. Parts I and II and *The Maryvale Project*. Buffalo, New York: Inter-University Project One at State University of New York at Buffalo, Cornell University, The University of Rochester, and Syracuse University with the support of the Ford Foundation, 1948.

Fischer, John H. Review of *Turmoil in Teaching* by T. M. Stinnett. *Phi Delta Kappan*, Vol. 50, No. 2, October, 1968.

Flanders, N. A. *Interaction Analysis in the Classroom: A Manual for Observers*. Unpublished manuscript. Ann Arbor: University of Michigan, 1960.

Gage, N. L. (ed.) *Handbook of Research on Teaching*. American Educational Research Association. Chicago: Rand McNally, 1963.

Gallagher, James J., Mary Jane Aschner, and William Jenne. *Productive Thinking of Gifted Children in Classroom Interaction*. CEC Research Monograph Series B., No. B-5. Washington, D. C.: The Council for Exceptional Children, NEA, 1967.

Gerbner, George. "A Theory of Communication and its Implication for Teaching," *The Nature of Teaching*. Louise M. Berman. (ed.) Milwaukee, Wisc: School of Education, University of Wisconsin at Milwaukee, 1962.

Gleser, G. C., L. J. Cronbach, and N. Rajaratnam. "Generalizability of Scores Influenced by Multiple Sources of Variance," *Psychometriker*. Vol. 30, 1965.

Guilford, J. P. "Three Faces of Intellect," *American Psychologist*. Vol. 14, 1959.

Hatch, Winslow R. "What Standards Do We Raise?" *New Dimensions in Higher Education*. No. 12. U.S.Office of Education, 1963.

Hatch, Winslow R. "The Socratic Method in Modern Dress." *Improving College and University Teaching*. Graduate School-Oregon State University, Corvallis, Oregon, Summer 1957.

Hatch, Winslow R. "Inquiry Into Inquiry." *Improving College and University Teaching*. Graduate School-Oregon State University, Corvallis, Oregon, Autumn 1957.

Henry, Marvin A. *The Relationship of Difficulties of Teachers to Selected Aspects of the Professional Sequence of Education*. Unpublished doctoral dissertation. Bloomington, Indiana: Indiana University, June, 1963.

Hickey, Albert E., and John M. Newton. *The Logical Basis of Teaching: I. The Effect of Subconcept Sequence on Learning*. Newburyport, Mass.: ENTELEK, 1964.

Hollingsworth, H. L. *Judging Human Character*. Appleton-Century, 1922.

Hughes, Marie. *Development of the Means for the Assessment of the Quality of Teaching in the Elementary Schools*. Salt Lake City: University of Utah Press, 1959.

Hunt, Maurice P., and Lawrence E. Metcalf. *Teaching High School Social Studies*. New York: Harper and Bros., 1955.

Jenkins, William A. (ed.) *The Nature of Knowledge*. Milwaukee, Wisc.: School of Education, University of Wisconsin at Milwaukee, 1961.

LaGrone, Herbert. *A Proposal for the Revision of the Pre-Service Professional Component of a Program of Teacher Education*. Washington, D. C.: American Association of Colleges for Teacher Education, 1964.

Lantz, Donald L. "The Relationship of University Supervisors and Supervising Teachers' Ratings to Observed Student Teachers' Behavior," *American Educational Research Journal*, Vol. 4, No. 3, May, 1967.

Lindsey, Margaret. *New Horizons for the Teaching Profession*. Washington, D. C.: National Commission on Teacher Education, NEA, 1961.

Mars, Walter Jerry. *A Study Comparing Two Ways of Training Students to Observe the Classroom Teacher's Verbal Behavior*. Doctoral dissertation. Syracuse, New York: Syracuse University, 1964.

Marsh, Joseph E., and Eleanor W. Wilder. Identifying the Effective Instructor: A Review of the Quantitative Studies, 1900-1952. Air Force Personnel and Training Research Center, Lockland Air Force Base, San Antonio, Texas, 1958.

Medley, Donald M. "Coding Behavior with OScAR 4V," Pre-publication draft. Princeton, N. J.: Educational Testing Service (Mimeographed.)

Medley, Donald M. "The Language of Teacher Behavior: Communicating the Results of Structured Observations to Teachers," Paper presented at Annual Meeting of American Education Research Association, February, 1967. (Mimeographed.)

Medley, Donald M., Joseph T. Impellittere, and Lou H. Smith. "Coding Teachers Verbal Behavior in the Classroom," A manual for users of OScAR 4V. Copy provided by Medley. (Mimeographed.)

Medley, Donald M., and Donald L. Lantz. "Classroom Application of Teacher Behavior Research," Paper presented at the Annual Meeting of the American Education Research Association, February, 1967. (Mimeographed.)

Mitzel, H. E. A Behavioral Approach to the Assessment of Teacher Effectiveness. New York: Division of Teacher Education, College of the City of New York, 1957. (Mimeographed.)

Nuthall, G. A., and P. J. Lawrence. Thinking in the Classroom. New Zealand: New Zealand Council for Educational Research, 1965.

Phenix, Philip H. "The Architectonics of Knowledge," Education and the Structure of Knowledge. Stanley M. Elam (ed.). Chicago: Rand McNally, 1965.

Raths, Louis E., Selma Wassermann, Arthur Jonas, and Arnold M. Rothstein. Teaching for Thinking--Theory and Application. Columbus, Ohio: Charles E. Merrill, 1967.

Ryans, David. Characteristics of Teachers: Their Description, Comparison, and Appraisal. Washington, D. C.: American Council on Education, 1960.

Schueler, H., M. U. Gold, and H. E. Mitzel. The Use of Television for Improving Teacher Training and for Improving Measures of Student-Teaching Performance. Phase I--Improvement of Student Teaching. U. S. Office of Education, Govt. No. 730035. New York: Hunter College of the City University of New York, 1962. (Mimeographed.)

Simon, Anita, and E. Gil. Boyer. (eds.) Mirrors for Behavior--An Anthology of Classroom Observation Instruments. Under the Provisions of Title IV ESEA of 1965 in cooperation with the USOE Research Contract OEC 1-7-062867-3053. Philadelphia: Research for Better Schools, 1967.

Smith, B. O., Mary Jane Aschner, and Milton Meux. *A Study of the Logic of Teaching*. A report of the first phase of a five years' research project. *The Logical Structure of Teaching and the Development of Critical Thinking*. Project #258 (7257) U. S. Office of Education. Urbana, Illinois: Bureau of Educational Research, College of Education, University of Illinois. (Dittoed.)

Smith, B. Othanel, and Robert H. Ennis. *Language and Concepts in Education*. Chicago: Rand McNally, 1961.

Smith, B. Othanel, Milton Meux, Jerrold Coombs, Graham Nuthall, and Robert Precians. *A Study of the Strategies of Teaching*. USOE Cooperative Research Project Number 1640. Urbana, Illinois: Bureau of Educational Research, University of Illinois, 1967.

Smith, E. R., and R. W. Tyler. *Appraising and Recording Student Progress*. New York: Harper, 1942.

Taba, H., S. Levine, and F. F. Elzy. *Thinking in Elementary School Children*. Cooperative Research Project No. 1574, San Francisco: San Francisco State College, 1964.

Tibbetts, A. M. "To Encourage Reason on the Campus: A Proposal for a New College Course in Thinking and Writing." *American Association of University Professors Bulletin*, 54.4. Dec. 1968.

Turner, Richard L. *Problem Solving Proficiency Among Elementary School Teachers*. USOE Cooperative Research Project No. 1262. Bloomington, Indiana: Institute of Educational Research, 1964.

Wilk, Roger E., and William H. Edson. "A Study of the Relationships between Observed Classroom Behaviors of Elementary Student Teachers. Predictors of Those Behaviors, and Ratings by Supervisors," College of Education, University of Minnesota, January, 1962.

Wilson, John. *Language and the Pursuit of Truth*. Cambridge, England: University of Cambridge Press, 1960.

Winer, B. J. *Statistical Principles in Experimental Design*. McGraw-Hill: New York, 1962.

Wittgenstein, Ludwig. *Philosophical Investigations*. Oxford: Basil Blackwell, 1958.

Woodruff, Asahel D. *Basic Concepts of Teaching*. San Francisco: Chandler, 1961.

Yee, Albert H. *The Student Teacher Trend*. Research Report to the U. S. Office of Education. Austin, Tex.: The University of Texas, 1967.

APPENDIX A

GUIDE FOR ASSESSMENT OF SECONDARY STUDENT TEACHER PERFORMANCE

Procedures for Use of the Secondary Student Teacher Performance Profile (SSTPP) in Observing, Rating, Recording, and Reporting

I. Supervisor's Observation Procedure

A. Preparation for Observation

1. The college supervisor and student teacher should become acquainted at a seminar immediately preceding student teaching.

2. An administrative visit is to be made to the student in his school during the first two weeks of student teaching. This visit provides for:

    a. Orientation of principal and supervising teacher to the student teaching program, and

    b. A brief conference with student teacher concerning schedule, adjustment, problems, etc.

    The SSTPP form is not used on the first visit.

3. The college supervisor should be familiar with:

    a. "Personal Information Form"--A four-page document completed by the student when he applies for student teaching, and

    b. Record of courses taken and grades earned in teaching field, and

    c. The Student Teaching Journal kept by student and supervising teacher. The Journal provides a log of the student's activities after arriving at the school along with a cooperative section which includes exchanges between the student teacher and his supervising teacher concerning philosophical and procedural questions and matters of detail. It also contains the student teacher's fortnightly summary.

B. Observation Procedure

1. The supervisor should greet student teacher and supervising teacher prior to class.

2. Observations should be made from an inconspicuous position in the classroom.

C. Rating and Recording Procedures

1. Ratings should be recorded immediately after the end of the observation.

2. All items for which evidence is observed are to be marked by an X on the number deemed appropriate. Where both poles of the criterion are approximately equally valid the middle number 4 should be marked.

3. Items for which no evidence is observed are to be marked on the N.

4. Questions on critical incidents affecting interpretation of a particular quality are to be written on the back of the card.

5. Additional rating procedures to be followed during or immediately following class:

   a. A check (✓) may be placed on items numbered II and V in addition to an X if any insight or information is received from sources other than the classroom visit which contradicts the rating made during the classroom visit.

   b. The following data are to be recorded on the card:
      (1) week of student teaching in which observation was made
      (2) grade level of students in class
      (3) subject being taught
      (4) date of observation
      (5) whether it was the second or third visit
      (6) which experimental treatment was being applied.

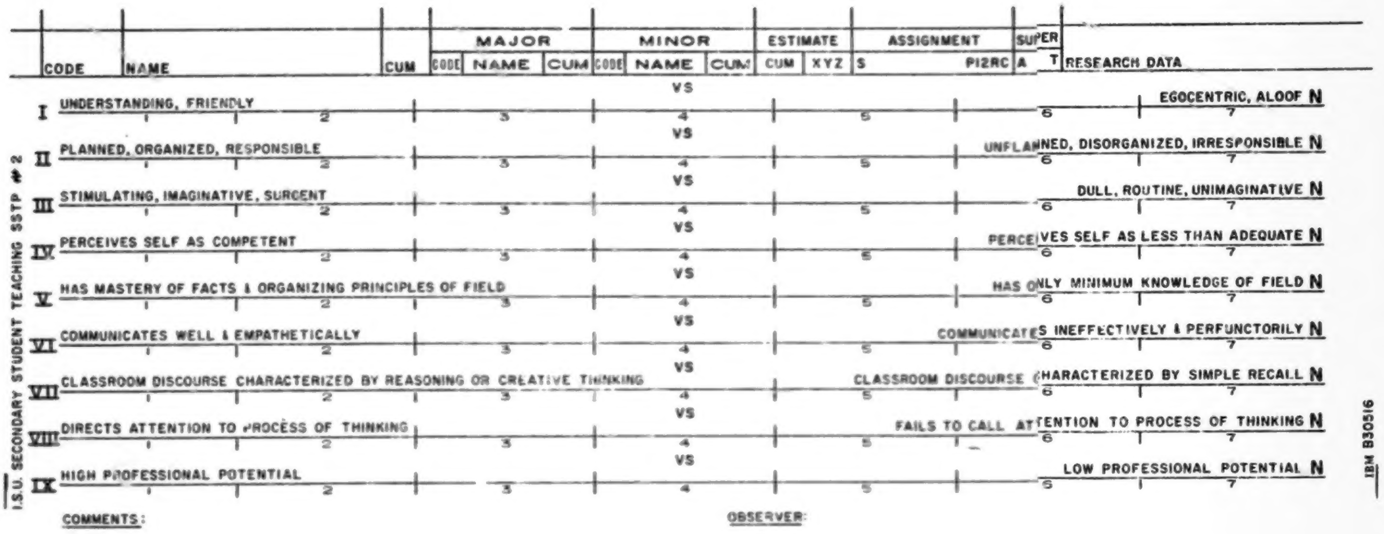
Completed cards are to be turned in to director at the end of the week.

II. General Policies

College supervisors do not have access to previous ratings of student teachers when making further observations.

No explanation of the SSTPP should be given to the student teacher or the supervising teacher at this time. After the reliability and validity have been determined we may decide to discuss the observations with the persons involved.

III. SSTPP Form 2



I.S.U. SECONDARY STUDENT TEACHING SSTP # 2

			MAJOR			MINOR			ESTIMATE		ASSIGNMENT		SUPER		
CODE	NAME	CUM	CODE	NAME	CUM	CODE	NAME	CUM	CUM	XYZ	S	P12RC	A	T	RESEARCH DATA

I	UNDERSTANDING, FRIENDLY	1	2	3	VS 4	5	6	7	EGOCENTRIC, ALOOF N
II	PLANNED, ORGANIZED, RESPONSIBLE	1	2	3	VS 4	5	6	7	UNPLANNED, DISORGANIZED, IRRESPONSIBLE N
III	STIMULATING, IMAGINATIVE, SURGENT	1	2	3	VS 4	5	6	7	DULL, ROUTINE, UNIMAGINATIVE N
IV	PERCEIVES SELF AS COMPETENT	1	2	3	VS 4	5	6	7	PERCEIVES SELF AS LESS THAN ADEQUATE N
V	HAS MASTERY OF FACTS & ORGANIZING PRINCIPLES OF FIELD	1	2	3	VS 4	5	6	7	HAS ONLY MINIMUM KNOWLEDGE OF FIELD N
VI	COMMUNICATES WELL & EMPATHETICALLY	1	2	3	VS 4	5	6	7	COMMUNICATES INEFFECTIVELY & PERFUNCTORILY N
VII	CLASSROOM DISCOURSE CHARACTERIZED BY REASONING OR CREATIVE THINKING	1	2	3	VS 4	5	6	7	CLASSROOM DISCOURSE CHARACTERIZED BY SIMPLE RECALL N
VIII	DIRECTS ATTENTION TO PROCESS OF THINKING	1	2	3	VS 4	5	6	7	FAILS TO CALL ATTENTION TO PROCESS OF THINKING N
IX	HIGH PROFESSIONAL POTENTIAL	1	2	3	VS 4	5	6	7	LOW PROFESSIONAL POTENTIAL N

COMMENTS:

OBSERVER:

IBM B30516

July 2, 1968

REVISED INSTRUCTIONS FOR CRITERION VII (EXPERIMENTAL)

Score as a "1"

If the major share of teacher's questions are open-ended or thought-provoking.

If teacher frequently translates principles into terms which are in the pupils' fields of experience.

Score as a "4"

If teacher ignores good opportunities for thought part of the time and seizes them part of the time.

Score as a "7"

If students are never asked "why?" or "how?"

If students are not permitted a moment to think of answer.

If no analogies to students' experiences are given.

REVISED INSTRUCTIONS FOR CRITERION VIII (EXPERIMENTAL)

Score as a "1"

If teacher makes students aware of the problem and the process for solving it.

If students actually comprehend the problem.

Score as a "4"

If teacher understands the problem and meaning of terms but fails to communicate it to students.

If teacher assumes students will comprehend the process.

Score as a "7"

If teacher never identifies the problem nor defines terms.

GLOSSARY OF OPERATIONAL DEFINITIONS FOR SECONDARY STUDENT TEACHER PERFORMANCE PROFILE

Criterion I--Understanding, Friendly vs. Egocentric, Aloof

Teachers who are described as having a high degree of the quality indicated by this criterion are characterized by such terms as friendly, understanding, tactful, good-natured, sympathetic, kindly, democratic, fair, approachable, gregarious, cooperative, genial, "looks on the bright side," enjoys the contact with his class. Superficial friendliness and mothering behavior should be recognized and given a low rating.

Criteria I, II, and III are virtually Ryans' XYZ factors.

Illustrative classroom behaviors: (The specific behaviors are considered to be parallel or synonymous. Students need not rate high on all examples to be rated high.)

Understanding, Friendly 1 2 3	*vs.* 4	*Egocentric, Aloof* 5 6 7
1. Shows concern for a pupil's personal emotional problems and needs.		1. Recognizes only academic accomplishments of pupils; no concern for personal problems; scolds.
2. Tolerant of errors on part of pupils.		2. Completely unsympathetic with a pupil's failure at a task.
3. Finds good things in pupils to which to call attention.		3. Calls attention only to very poor work.
4. Listens encouragingly to pupil's viewpoint.		4. Impatient.
5. Behavior that invites friendliness from all pupils.		5. Stiff and formal in relations with pupils.
6. Goes out of way to be pleasant and/or to help pupils.		6. Hypercritical, fault finding, apart, removed from pupil activity.
7. Smiles and speaks to individual pupils; knows names.		7. Reserved and sober; doesn't know names of students.

Basic References:

David Ryans, *Characteristics of Teachers: Their Description, Comparison, and Appraisal* (Washington, D.C.: American Council on Education, 1960).

E. J. Amidon and Ned A. Flanders, *The Role of the Teacher in the Classroom* (Minneapolis: Paul S. Amidon & Associates, Inc., 1963).

Criterion II--Planned, Organized Responsible vs. *Unplanned, Disorganized Irresponsible*

Teachers associated with a high degree of the quality indicated by this criterion are described as businesslike, systematic, responsible, consistent, definite, thorough, self-controlled, well-prepared.

Illustrative classroom behaviors:

Planned, Organized, Responsible vs. *Unplanned, Disorganized, Irresponsible*

1 2 3 4 5 6 7

1. Has a detailed lesson plan; has materials that will be used.	1. Has no lesson plan; materials are not available when he needs them.
2. Tells class what to expect during period; has next assignment planned.	2. Class members see no organization in what they are doing; next assignment unplanned.
3. Objectives of the lesson are clearly discernible; proceeds with assurance.	3. Objectives not apparent; undecided as to next step.
4. Anticipates and prepares answers for student questions.	4. Is surprised and confused by questions.
5. Keeps discussion focused on objectives.	5. Easily distracted from matter at hand; course of action easily swayed by circumstances of the moment.
6. Is aware of school regulations and is guided by them.	6. Ignores school regulations.
7. Punctual.	7. Does not perform own duties on time.
8. Gives definite directions and checks on responses.	8. Assignments and directions indefinite; no evidence of follow-up.

To be Checked Out of Class

1. Does he understand and is he following directions given in seminar?	1. Vague about assignments and is not doing them well.
2. Meets obligations in record keeping.	2. Careless and evasive in record keeping.

Basic reference:
David Ryans, Characteristics of Teachers: Their Description, Comparison, and Appraisal (Washington, D.C.: American Council on Education, 1960).

Criterion III--Stimulating, Imaginative vs. Dull, Routine
Surgent Unimaginative

Teachers who possess high degrees of this quality are described as original, adaptable, stimulating, resourceful, imaginative, "puts pupils on their own," encourages pupil initiative. This is the Z factor described in Ryans.

Illustrative classroom behaviors:

Stimulating,	*Imaginative,*	*Surgent*	*vs.*	*Dull,*	*Routine,*	*Unimaginative*
1	*2*	*3*	*4*	*5*	*6*	*7*

1. Highly interesting presentation; got and held attention.	1. Class seems bored, uninterested.
2. Clever and witty; enthusiastic, animated.	2. Pedantic, boring, phlegmatic, bored-acting.
3. Assignment challenging.	3. Fails to provide challenge.
4. Took advantage of pupil interest.	4. Failed to capitalize on pupil interest.
5. Seemed to provoke thinking: provoked strong esthetic appeal.	5. Class lacked enthusiasm; lacked interest.
6. Uses what seem to be original and relatively unique devices to aid instruction.	6. Uses routine procedures without variation.
7. Tries new materials or methods.	7. Does not depart from procedure to take advantage of a relevant question and situation.
8. Resourceful in answering questions; has many pertinent and novel illustrations available.	8. Not resourceful in answering questions or providing explanation.

Basic reference:

David Ryans, Characteristics of Teachers: Their Description, Comparison, and Appraisal (Washington, D.C.: American Council on Education, 1960).

Criterion IV--Perceives Self as Competent vs. Perceives Self as Less Than Adequate

Teachers who possess a high degree of this quality tend to see themselves as liked, wanted, worthy, and able to do a good job of whatever they undertake. This is an attempt to rate self-confidence.

Illustrative classroom behaviors:

Perceives Self as Competent			*vs.*	*Perceives Self as Less Than Adequate*		
1	*2*	*3*	*4*	*5*	*6*	*7*

1. Speaks confidently; self-confident in relations with pupils.	1. Faltering, hesitant in speech, timid in relation with pupils.
2. Takes mistakes and/or criticism in stride.	2. Disturbed and embarrassed by mistakes and/or criticism.
3. Actions are reasonable, aggressive, and spontaneous.	3. Appears timid and shy.
4. Poised.	4. Lacks poise.
5. Accepts new tasks readily.	5. Offers excuses for not acting, delays.

Basic references:

Arthur W. Combs (ed.), Perceiving, Behaving, Becoming, Yearbook, 1962 (Washington, D.C.: Association for Supervision and Curriculum Development, 1962).

Arthur W. Combs, The Professional Education of Teachers (Boston: Allyn & Bacon, 1965).

Robert E. Bills, About People and Teaching, Bulletin of the Bureau of School Service, Vol. XXVIII, No. 2 (Lexington: University of Kentucky, December, 1955).

Criterion V--Has Mastery of Facts and Organizing Principles of Field vs. Has Only Minimum Knowledge of Field

A teacher who possesses a high degree of this quality recognizes important and significant knowledge in his field--concepts, generalizations, and relationships. He recognizes the organizing principles, or the basic ideas, of how things are related. This is sometimes referred to as the inquiry structure of the discipline.

Illustrative classroom behaviors:

Has Mastery of Facts and Organizing Principles of Field	*vs.*	*Has Only Minimum Knowledge of Field*
1 2 3	*4*	*5 6 7*
1. Presentation concerned with basic principles of the field.		1. Presentation simply reviews the facts from the text.
2. Organized ideas, concepts, and principles on a recognizable priority system.		2. Presentation or discussion was not directed toward any central idea or generally accepted generalizations.
3. Drew examples and explanations from various sources and related fields.		3. Failed to enrich the class with illustrations familiar to students or from related fields.
4. Leads his class to form and test hypotheses.		4. Inaccurate or irrelevant comments; emphasizes mastery of unrelated facts.
5. Approach to subject was in spirit of discovery.		5. Approach to subject was one of routine examination of a given amount of materials.

Basic references:

Harry S. Broudy, B. O. Smith, and J. R. Burnett, Democracy and Excellence in American Secondary Education (Chicago: Rand McNally, 1964).

Jerome Bruner, The Process of Education (Cambridge, Mass.: Harvard University Press, 1962).

Albert E. Hickey and John M. Newton, The Logical Basis of Teaching: I. The Effect of Subconcept Sequence on Learning (Newburyport, Mass.: ENTELEK Incorporated, 1964).

Basic references (Continued):

William A. Jenkins (ed.), The Nature of Knowledge (Milwaukee, Wisconsin School of Education, University of Wisconsin at Milwaukee, 1961).

Philip H. Phenix, "The Architectonics of Knowledge," in Stanley M. Elam (ed.), Education and the Structure of Knowledge (Chicago: Rand McNally, 1965).

Criterion VI--Communicates Well and Empathetically vs. *Communicates Ineffectively and Perfunctorily*

Teachers who possess a high degree of this quality show acute sensitivity to the perceptions of pupils and make use of these insights to make presentations at their level of understanding. A detailed analysis of the communication function is found in Smith and Ennis.

Illustrative classroom behaviors:

Communicates Well and Empathetically	*vs.*	*Communicates Ineffectively and Perfunctorily*
1 2 3	*4*	*5 6 7*
1. Reacts to expression of students by asking if there are questions or whether a student can give an example.		1. Presents material without recognizing obvious indications of confusing or personal interest on part of individuals.
2. Draws examples from local community or current interests of youth of the age group being taught.		2. Uses few examples or ones that serve to confuse the student.
3. Expresses pleasure at receiving student questions and comments.		3. Shows impatience with student interruptions and digressions.
4. Makes effective use of audio-visual aids to supplement communication.		4. Makes no attempt to adapt activities or materials to needs and interests of students.
5. Analyzes and answers student questions efficiently.		5. Makes little attempt to understand question; answers in rambling fashion.
6. Has no distracting mannerisms.		6. Has nervous mannerisms which distract.
7. Speaks fluently.		7. Speaks hesitantly.
8. Diagnoses readiness of students.		8. Unaware of student interest and understanding.

Basic references:

George Gerbner, "A Theory of Communication and its Implication for Teaching," in Louise M. Berman (ed.), The Nature of Teaching (Milwaukee, Wis.: School of Education, University of Wisconsin at Milwaukee, 1962).

B. O. Smith and Robert H. Ennis, Language and Concepts in Education, Chapter 6 (Chicago: Rand McNally, 1961).

Criterion VII--Classroom Discourse Characterized by Reasoning and Creative Thinking vs. Classroom Discourse Characterized by Simple Recall

A teacher who possesses a high degree of this quality helps students to go beyond specific recall of facts into an understanding and application of the problem-solving process.

Illustrative classroom behaviors:

Classroom Discourse Characterized by Reasoning and Creative Thinking	*vs.*	*Classroom Discourse Characterized by Simple Recall*
1 2 3	*4*	*5 6 7*
1. Seeks definition of problems with his class.		1. Raises own questions and answers them; or uses questions given and answered in the text.
2. Leads his class in a search for possible solutions to problems.		2. Assumes a single correct solution to a problem.
3. Leads class to decide upon the most promising solution to problems.		3. Gives the single "best" solution.
4. Leads class to evaluate and revise solutions previously reached by the class.		4. Requires uncritical acceptance of facts.
5. Encourages students to make wider application of general principles discovered.		5. Teaches facts unrelated to application.
6. Major share of teacher's questions are open-ended.		6. Students are never asked, "Why?"
7. Permits a time lapse to occur after asking a question.		7. Students seldom given even a moment to think.
8. Frequently relates to experiences of pupils.		8. Makes no analogies to students' experiences.

Basic references:

Mary J. Aschner, "The Analysis of Classroom Discourse--A Method and Its Uses" (unpublished doctoral dissertation, University of Illinois, Urbana, 1958).

John C. Wilson, Language and the Pursuit of Truth (Cambridge, England: Cambridge Press, 1960).

Category VIII--Directs Attention to the Logical Operations in Thinking vs. Fails to Call Attention to the Logical Operations in Thinking

A teacher who finds opportunities to point out and analyze the logical operations in thinking will score high on this criterion Teachers who help pupils become aware of the processes they follow in identifying and solving problems will score high. Teachers who are concerned only with subject matter will score low.

Illustrative classroom behaviors:

Directs Attention to the Logical Operations in Thinking vs. *Fails to Call Attention to the Logical Operations in Thinking*

1 2 3 4 5 6 7

Directs Attention to the Logical Operations in Thinking	*Fails to Call Attention to the Logical Operations in Thinking*
1. Recognizes and points out that confusion arises from different definitions of terms. "You can't disagree unless you are talking about the same thing."	1. Defines terms only according to what the dictionary says. Uses authoritarian rather than agreed upon meanings as basis for definition.
2. Recognizes and points out that differences or arguments arise from contrasting views of what is worthwhile or best.	2. Ends an argument with a shrug and such comment as: "That's one person's opinion. I have mine. Let's get back to the subject."
3. Points out the differences between what is observed and what is inferred from the observation.	3. Fails to distinguish between fact and inference.
4. Asks for and demands examination of evidence.	4. Makes frequent dogmatic statements.
5. Faces up to the problem where conflicting evidence or contradictory facts are presented by teaching the laws of evidence and rules of proof.	5. Ignores cases of loose reasoning or contradictory statements.
6. Leads students to state assumptions.	6. Asks a question that requires thinking but doesn't wait for students to think.
7. Helps students find their way through the steps of problem solving.	7. Either gives answer or avoids the problem.
8. Makes beliefs truer.	8. Accepts beliefs and opinions uncritically.

Basic references:

Arno A. Bellack, Ronald T. Hyman, Frank L. Smith, Jr., and Herbert Kliebard, *The Language of the Classroom*, Meanings Communicated in High School Teaching, Part II, USOE Cooperative Research Project No. 2023, 1965.

Kenneth D. Benne, George E. Axtelle, B. Othanel Smith, and Bruce Raup, *The Improvement of Practical Intelligence* (New York: Harper, 1950).

Morris R. Cohen and Ernest Nagel, *An Introduction to Logic and the Scientific Method* (New York: Harcourt, Brace, 1934).

James J. Gallagher, Mary Jane Aschner, and William Jenne, *Productive Thinking of Gifted Children in Classroom Interaction*. CEC Research Monograph Series B., No. B-5 (Washington, D.C.: The Council for Exceptional Children, NEA, 1967).

Louis E. Raths, Selma Wassermann, Arthur Jonas, and Arnold M. Rothstein, *Teaching for Thinking--Theory and Application* (Columbus, Ohio: Charles E. Merrill, 1967).

B. O. Smith and Milton Meux, *A Study of the Logic of Thinking*, USOE Cooperative Research Project No. 258 (7257) (Urbana, Ill.: Bureau of Educational Research, University of Illinois, 1965). (Trial Edition.)

H. Taba, S. Levine, and F. F. Elzey, *Thinking in Elementary School Children*, USOE Cooperative Research Project No. 1574 (San Francisco State College, 1964).

John C. Wilson, *Language and the Pursuit of Truth* (Cambridge, England: Cambridge Press, 1960).

Asahel D. Woodruff, *Basic Concepts of Teaching* (San Francisco: Chandler, 1961).

Criterion IX--High Professional Potential vs. *Low Professional Potential*

The rating on this criterion is not an average of the other qualities in this profile. It goes beyond these qualities to include all factors and feeling concerning the professional potential of the student teacher. This includes such considerations as success in a typical school situation, desire to achieve the maximum potential for himself and his students, and interest in the improvement of the profession. It constitutes a global estimate.

Basic references:

Louise M. Berman (ed.), The Nature of Teaching (Milwaukee, Wis.: School of Education, University of Wisconsin at Milwaukee, 1962).

J. Bruce Biddle and William J. Ellena (eds.), Contemporary Research on Teacher Effectiveness (New York: Holt, Rinehart & Winston, 1964).

Arthur W. Combs, The Professional Education of Teachers (Boston: Allyn & Bacon, 1965).

APPENDIX B

INSTRUCTIONS FOR USING THE TEACHER CLASSROOM ACTIVITY PROFILE (TCAP) AND DEFINITIONS AND ILLUSTRATIONS OF TERMS

The Teacher Classroom Activity Profile (TCAP) is printed in two forms. The 8" x 11" no-carbon-required triplicate form B and the IBM card form C. The NCR form B will generally be used to record observations in the classroom--one copy (yellow) is for the student teacher, one copy (white) is for the Division of Teaching, and one copy (pink) is for the Departmental Supervisor. The IBM card will be used for key punching, filing, and research purposes.

The completed profile provides data which enable the student and supervisor to "reconstruct the experience" and examine the effectiveness of the teaching acts and consider alternatives. The completed profile constitutes a sequential account of the major activities in which the teacher engages during the class session.

Instructions for Use of TCAP

The observer records a continuous line moving among the seven major activities in three-minute intervals. Explanatory notes should be keyed to the column number which indicates the sequence of three-minute intervals. If there is just a momentary shift in categories, a vertical line going up or down to the proper category should be made without interruption of the general flow of the regular profile graphs.

It has been found helpful to indicate the time of day at each of the three-minute intervals, starting in column 1 with the minute the class starts and then recording the time at three-minute intervals after that in the numbered squares. If a teacher interrupts a presentation to reprimand a student or to ask for attention, since this activity is classified as Management--Non-Learning, a line would go up to section MN or, similarly, if the teacher is conducting a recitation and stops a moment to ask a question which really provokes thinking and then goes on with the recitation, a line would drop down to Logical Thinking. If this activity goes on for a minute or more, the graph should show that as a part of the continuous line. It is possible to indicate one- or two-minute sequences by using one-third or two-thirds of the space. For summary purposes, count four such lines as the equivalent of one minute. TCAP forms should be completed including the summary computation and returned to Research Office of the Division of Teaching at the end of each platoon. Be sure to include identification data--student and observer.

The total number of minutes should, of course, add up to the total number of minutes spent in observation. The percentage should be computed approximately--that is, to the nearest percent. Percentages should tally to 100 percent. A table to facilitate this computation is available in Dr. Beasley's office. The NCR paper makes it possible for the observer to provide students with a copy of the analysis. The left-hand column, "explanatory notes," should be used to describe any unusual movement of the profile or to identify something that happened that might be discussed. The right-hand section, called "anecdotal records," could be used to make evaluative judgments, jot down ideas, or make suggestions that may be helpful to the student. Some of the staff have found it helpful to mark an X in the appropriate category at the time an opportunity to help students improve their thinking was missed.

See attached example of completed form.

Definitions of Major Categories of TCAP

While the terms used to identify the seven categories of teacher activity carry common connotations, their use in this study is restricted to the precise meanings as defined and illustrated below. The definitions and examples serve as a basic point of reference. The precise distinctions among the various categories are clarified in staff discussions and individual conferences. TCAP can be used in both academic classes and laboratory or shop-type classes; however, the precise meaning of some categories varies with the type of class.

Glossary of terms as applied to academic-type classes:

(English, Social Studies, Mathematics, Science other than Science Laboratories, etc.)

MN Management--Non-Learning

Management of classroom when the teacher is not attempting to teach, e.g., reading announcements, taking roll, distributing materials, organizing equipment, idle time, disciplining pupils, waiting for the bell to ring.

ML Management--Learning

Management of classroom so that learning may occur but the teacher is not involved except in a managerial role, e.g., showing a sound film, administering a written examination, supervising study time, student reports.

P Presentation

The presentation of subject matter by the teacher in some organized fashion, e.g., lectures, demonstrations, illustrated talks, blackboard presentation, reading.

R Recitation/Drill

The solicitation of student responses which call for terse memorized data, oral testing to determine if assignments have been read, review questions, etc. Include drill and practice time here.

D Discussion/Random

Random discussion involving student-teacher interaction but without analysis or synthesis. "Stream-of-consciousness" discussion without any apparent focus or purpose except to consume time until the period is over, e.g., "talk-talk-talk." When discussion does come to a point it fits into the next category, LT, Logical Thinking.

LT Logical Thinking

Discussion which involves analysis and synthesis. The teacher is deliberately encouraging or permitting thinking to occur. This category is more than reciting or repeating something which has been learned or memorized. When the teacher acts to encourage thinking, it should be recorded here whether or not the act is successful. Use vertical lines to this category when a brief interval of thinking occurs in presentations, recitations, or discussion.

TP Thinking Process

Deliberate, conscious attention on the part of the teacher to the intellectual process, e.g., pointing out to students the factual and/or logical basis of their thinking, pointing out errors in reasoning, examining the reliability and validity of evidence, examining the adequacy of the sample, defining terms, checking assumptions, examining the scientific method, examining values, seeking reason for conflicting opinions, locating the source of difficulty, examining the "method of inquiry." Use explanatory notes to report good examples. Put an X on the profile if an opportunity to contribute to the thinking process is missed.

Glossary of terms as applied to laboratory-type classes:

(Science Laboratory, Shop, Band, Orchestra, Physical Education, Typing, etc.)

MN Management--Non-Learning

Basically the same as for academic. Include clean-up and wasted time. Record here time in which teacher ignores pupils to work on outside activities.

ML Management--Learning

Basically the same as for academic. Warm-up time, showering, dressing, when supervised by teacher.

P Presentation

Same as for academic.

R Recitation/Drill

Same as for academic.

D Discussion/Random

Same as for academic.

LT Logical Thinking

Consider all purposeful work which the teacher is supervising or directing in this category.

TP Thinking Process

Same as for academic. Applies if teacher is working with a single teacher, a group, or the total class.

The model which evolved from the staff study identified four levels of intellectual activity which could be categorized from observations of teacher behavior.

Levels of Intellectual Activity	Teacher Activity
Level One: Managing	Directing, Commanding, Managing, Rebuking, Accepting Students. These are usually imperative statements.
Level Two. Informing	Presenting Information, Reporting, Describing, Recalling from Memory. Getting students to do these. These are usually empirical statements.
Level Three: Thinking	Summarizing, Classifying, Explaining, Reasoning and Evaluating. Getting students to do these. These are usually relational statements.
Level Four: Thinking Process	Thinking about Thinking; Analyzing the Processes of Thinking Going on in Others; Recognizing the Roles Played by Assumptions, Biases, and Values; Locating Sources of Misunderstanding. Getting students to do these. These are usually logical statements.

This analysis of thinking proved helpful in identifying the various categories for classifying student teacher behavior on the Teacher Classroom Activity Profile. It also provided the basis of making evaluative ratings on Criteria VII and VIII of the Secondary Student Teacher Performance Profile.

The schema which follows suggests relationships between this system for looking at the teacher's intellectual activities and the staff's interpretation of the way some other scholars have organized their thinking. The relationships are approximations and tangential. The purpose of showing them here is to give added meaning to this system of classification rather than to compare one with another.

RELATIONSHIPS BETWEEN THE SCHEMA DESIGNED FOR THIS STUDY AND THOSE DEVELOPED BY OTHERS FOR ORGANIZING IDEAS ABOUT THINKING AND THE TEACHING ROLE

Sharpe and Staff	Wilson[1]	Gallagher[2]	Raths[3]	Aschner[4]
Level One: Managing	Imperative and Attitude Statements	Routine		
Level Two: Informing	Empirical Statements	Cognitive Memory	Observing Collecting Data	Remembering
Level Three: Thinking	Empirical Statements	Convergent Thinking Divergent Thinking Evaluative Thinking	Summarizing Comparing Classifying Imagining Hypothesizing Applying Facts and Principles Making Decisions Designing Projects	Reasoning Evaluating or Judging Creative Thinking
Level Four: Thinking About Thinking	Analytical Statements		Looking for Assumptions	

---

[1] John Wilson, Language and the Pursuit of Truth (Cambridge, England: University of Cambridge Press, 1960).

[2] James J. Gallagher, Mary Jane Aschner, and William Jenne, Productive Thinking of Gifted Children in Classroom Interaction. CEC Research Monograph Series B., No. B-5 (Washington, D.C.: The Council for Exceptional Children, NEA, 1967).

[3] Louis E. Raths, Selma Wassermann, Arthur Jonas, and Arnold M. Rothstein, Teaching for Thinking--Theory and Application (Columbus, Ohio: Charles E. Merrill, 1967).

[4] Mary J. Aschner, The Analysis of Classroom Discourse--A Method and Its Uses, unpublished doctoral dissertation (Urbana, Ill.: University of Illinois, 1958).

RELATIONSHIPS BETWEEN THE SCHEMA DESIGNED FOR THIS STUDY AND THOSE DEVELOPED BY OTHERS FOR ORGANIZING IDEAS ABOUT THINKING AND THE TEACHING ROLE
(Continued)

Sharpe and Staff	Smith[5]	Medley[6]	Bloom[7]	Woodruff[8]
Level One: Managing	Directing and Managing	Affective Statements (Considering and Rebuking) Procedural Statements Directing and Describing		
Level Two: Informing	Describing Stating Reporting	Substantive Statements--Informing	Memory	Level A--Show
Level Three: Thinking	Defining Designating Substituting Evaluating Opining Classifying Comparing and Contrasting Conditional Inferring Explaining	Substantive Statements--Problem Structuring	Translation Interpretation Application Analysis Synthesis	Level B--Discuss Level C--Apply Level D--Guide Original Thinking
Level Four: Thinking About Thinking			Evaluation	Some Aspects of Guiding Original Thinking

---

[5] B. O. Smith, Mary Jane Aschner, and Milton Meux, A Study of the Logic of Teaching, Project #258 (7257) U.S.Office of Education (Urbana, Illinois: Bureau of Educational Research, College of Education, University of Illinois).

[6] Donald M. Medley, "Coding Behavior with OScAR 4V," (Mimeograph) Pre-Publication Draft (Princeton, N.J.: Educational Testing Service).

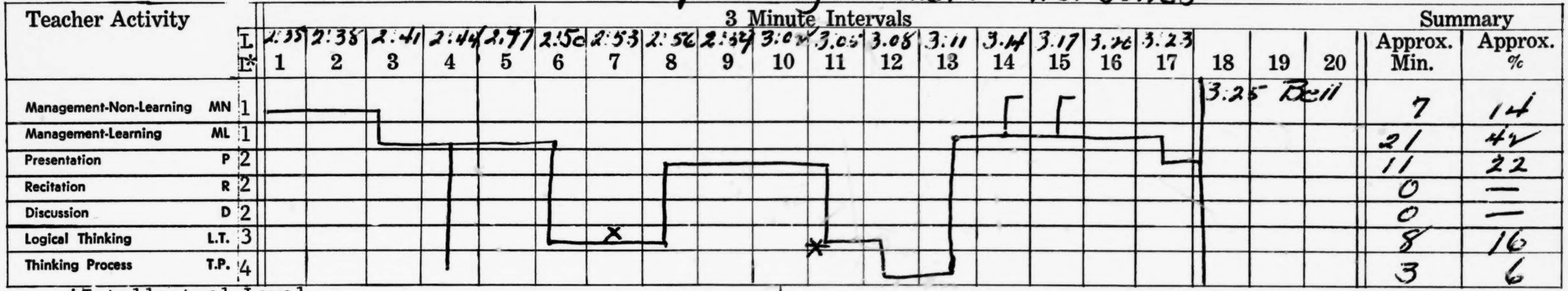
[7] Benjamin Bloom (ed.), Taxonomy of Educational Objectives (New York: Longmans, Green, 1958).

[8] Asahel D. Woodruff, Basic Concepts of Teaching (San Francisco: Chandler, 1961).

TEACHER CLASSROOM ACTIVITY PROFILE STUDENT TEACHER Sandra Blank Date 12-14-67
TCAP FORM B

I.S.U. DIVISION OF TEACHING, 1966 CLASS English -V TYPE Academic SUPERVISOR Sharpe

Supervising Teacher: Mrs. Jones



Teacher Activity	L I*	3 Minute Intervals																				Summary	
		2:35 1	2:38 2	2:41 3	2:44 4	2:47 5	2:50 6	2:53 7	2:56 8	2:59 9	3:02 10	3.05 11	3.08 12	3:11 13	3.14 14	3.17 15	3.20 16	3.23 17	18	19	20	Approx. Min.	Approx. %
Management-Non-Learning MN	1																		3:25 Bell			7	14
Management-Learning ML	1																					21	42
Presentation P	2																					11	22
Recitation R	2																					0	—
Discussion D	2																					0	—
Logical Thinking L.T.	3							x				*										8	16
Thinking Process T.P.	4																					3	6

*Intellectual Level

Explanatory Notes

-3 Distributed corrected themes—individually—without comments

3-6 Allowed time to read her comments.

7-8 Responded to questions about marks — Good.

8-11 Informal introduction to the "Short Story."

12 Discussion of values — "good and bad."

13-17 Study time

17 Two minutes for assignment

Anecdotal Records

* Good response to pupil who asked why he should study short stories.

4 Good question

7 Failed to clarify the definition

14-15 Reprimanded pupils

← Teacher played role of umpire

13-17 Teacher walked around room — Good

CUE CHART FOR CLASSIFYING TEACHER BEHAVIOR ON TCAP

TCAP CATEGORIES	LEVELS OF THINKING	KINDS OF STATEMENTS	TEACHER'S PRIMARY INTENTIONS OR CENTRAL PURPOSES	CLUE WORDS
Management--Non-Learning	Level One	Imperative--Affective	To control pupils To direct their behavior	Please (All commands) Class
Management--Learning (Originally included in Management)	Level One	Imperative--Affective	To influence classroom climate	Please (All commands)
Presentation	Level Two	Reporting--Stating--Recalling information from memory	To inform students	Now! Give me your attention Take notes on this
Recitation/Drill	Level Two	Questions and responses for getting students to report, state, recall information	To discover what pupils know or do not know To inform pupils	What - - - - ? How - - - - ?
Discussion/Random (Originally called Discussion I)	Levels One and Two	Affective--Responding without thinking-- Stream-of-consciousness talk	To kill time To impress students as contrasted with informing them To establish relationships	One time I - - - - Did you hear about - - ? Ah! Ah! - - - -
Logical Thinking (Originally called Discussion II)	Level Three	Questions and statements which help students to think, to summarize, to reason, to evaluate, to analyze, to discover	To get students to think To guide students to discovery and verification	Why? followed by pause How? followed by pause What is the evidence? What are the assumptions? What are the inferences?
Attention to Thinking Process	Level Four	Statements about language and logic which help students to analyze the ways in which they come to empirical conclusions and normative judgments-- Statements which help them locate sources of misunderstanding and disagreement	To provide students with the skills for thinking To inform them of accepted ways of using language To provide skills for pursuing truth	What do you mean? How do you define that? Is the sampling adequate? Where is our hang-up?

APPENDIX C

ANECDOTAL REPORTS
ILLUSTRATING CRITERION BEHAVIOR

The following précis, prepared by the staff of the Division of Teaching, were drawn from actual classroom observations to illustrate typical criterion estimates of teacher performance.

Only the relevant aspects of the encounters are reported. No attempt has been made to describe the complete context in which the event occurred. While none of these incidents by itself would determine the rating, it does provide an important clue to be weighed against other behaviors in determining the final rating.

These anecdotal reports when used in connection with the Glossary and behavioral definitions included in the "Guide for Assessment of Secondary Student Teacher Performance" should help observers make more reliable criterion ratings.

Three levels of anecdotal reports are provided for each criterion: those which illustrate a high rating of 1, those which illustrate a middle evaluation of 4, and those which illustrate a low rating of 7.

Criterion I--Understanding, Friendly vs. Egocentric, Aloof

Rating of 1--High

> In a seventh-grade, noisy, boisterous, low-track English class, Mary Jones was trying to gain control. On several occasions, the students asked questions at inopportune moments. When this occurred, Mary was quick to interrupt her presentation to answer the question. She was solicitous in her answers and seemed to be giving the student some inside information about how best to learn the material. As she introduced a new topic, she cautioned the students about pitfalls which they might encounter in learning the new topic. She assured the students that she was there for the purpose of making it easier for them to learn. After explaining a point, she asked five or six students, calling them by name, if they were sure that they understood. When a student answered correctly, she seemed to be delighted at the student's accomplishment. In conference after the class, the student teacher spoke favorably about various individuals in the class.

In a junior high English class, the student teacher announced that one of the students was moving from the community because his father had been transferred. She concluded with: "Bill, I am sure the class joins with me in wishing you the best of luck in your new school and don't forget to write to us about your new school."

Mr. Randolph was expressing his pleasure at the accomplishments of the students after they had exhibited to their parents their projects in science. "I was so proud of every one of you I just couldn't sleep! Timmy, wasn't it great your father could be home on leave? Wasn't he pleased with the thermometer you made?"

Rating of 4--Midpoint

In a typing class, the student teacher's directions were given to the entire class. She glanced impersonally at the students' work as she walked about the classroom. When a student raised his hand, she sometimes inquired, "Yes?" without using his name. She was patient in repeating directions. Occasionally she would smile at a student, but, for the most part, her facial expression was impassive. While her actions could not be rated as friendly, neither could they be considered unfriendly. At the end of the period one student remained to ask a question.

The students in a beginning instrumental music group were having trouble keeping together in playing the school song. "You do these parts OK when you play alone," Miss Gerda praised, but, "Now there is no earthly reason you can't stay with it when we put them together!"

The student teacher was observed in a junior high English class. During the first part of the period he smilingly spoke to the class, kidded some of the boys about their basketball game the previous night, asked a girl how long her arm would have to remain in a cast, and, in general, gave the impression that he was a good-natured, friendly person. His behavior was not consistent throughout the period, however. After collecting the weekly themes he noticed that one girl had not turned in hers and demanded why she had not prepared it. She informed him that she had written it at home the previous evening but that the school bus had come ten minutes early and that in her hurry to board the bus she had forgotten her theme. The student teacher said that this was no excuse and gave the entire class a short lecture on how they should grow up and learn to meet their responsibilities.

Rating of 7--Low

The student teacher was observed in a third-year high school German language class. She addressed them by their last names, using the German equivalent of Mr. and Miss. She was very reserved and sober before the class and created a "mightier than thou" image. At no time did she exhibit behavior that invited friendliness from her pupils. She missed several good opportunities to praise pupils for good work but called attention to poor work on the part of several students. She was impatient if a pupil did not respond rapidly and appeared unsympathetic when a pupil failed to give the correct response.

In a general science class, the student teacher gave a "quiz" during the first part of the period, had students to exchange papers, and read the answers aloud. Some of the slower students were subjected to ridicule as other students read aloud their greatly incorrect answers. Student teacher did nothing to correct this situation and indicated to me later "it is a fast way of grading papers."

It was a laboratory in biology. The project was to dissect the frogs. Annette was unable to proceed with the project. As Mr. Stone questioned her about her inactivity she said, "I honestly don't feel well." Mr. Stone's reply was, "Look, you'll never learn how to do this by being silly about it."

Criterion II--Planned, Organized, Responsible vs. Unplanned, Disorganized

Rating of 1--High

The unit Mr. Evans was teaching in chemistry involved the Periodic Table. He began his lesson saying, "During these past six weeks we have at times mentioned various compounds and elements without going into much detail on them. The periodic chart on the wall and the print-out I have given you will be very valuable in predicting the properties and reactivity of all the known elements. We will see the films

Chemical Families and Metals vs. Non-Metals. We also have models of atoms and their structure. From these you will see how and why elements react to form the many compounds that exist today."

In a general math class the student teacher's objective for the week was to teach students how to figure their families' gas, electric, and water expenses. The student teacher had meters which he had borrowed from the companies and was able to show students first-hand how to read the meter as well as how to compute the bill.

The student teacher quickly completed the attendance check and then turned to the distribution of papers to the students. The papers were arranged so that she could hand a group of papers to the students in the front seats and they could in turn pass back the correct paper to each student.

"Today we will begin by reviewing the quiz you took yesterday. We will take a look at each problem and discuss any points that need to be cleared up. After that we will take a look at the next set of problems and learn how to do those. After that, you will have about fifteen minutes to work on the homework so you can ask me about anything you do not understand. Friday we will finish this chapter and the test will be next Monday."

Rating of 4--Midpoint

The student teacher in health announced that the test would be postponed until the next day so that all the classes could take it on the same day. He then showed a film on the heart. After the film was over, he discussed it briefly. Following this, he reviewed the breathing process with a series of questions which traced air movement through the body. He then returned to a further discussion of the heart. He kept discussion focused on the objectives (in this case an impending test) and he asked questions in a logical sequence. He had not previewed the film in advance and he jumped from one topic to the next; then back to the original topic.

The student teacher in Latin II class was preparing his lesson for the next period when I found him. When I asked him if I were interfering with his preparation, he answered in the negative. After a few minutes of chatting, I asked him for his journal. He exclaimed, "Oh Yes!" and found it after a few moments of searching. When I asked him about his plans for the ensuing period, he stated that he would start with a certain list of exercises and that he hoped to finish the list during the period. The supervising teacher stated that the student teacher's lesson plans had been satisfactory. At the close of the period, the student had no suggestion about where to hold a conference.

After observing a sociology class that the student teacher had been teaching for four days I mentioned that I had noticed he was having difficulty pacing the class and had not gotten to the point in the material where his plans indicated he would be. He stated that he was having trouble developing adequate plans because, "This class likes to talk and I haven't yet found a way to shut them up when I want to move on."

Rating of 7--Low

In a health class that met only two days a week, a student asked the student teacher when he would be able to present the report he had been assigned four days prior. The student teacher replied that the class couldn't take time to listen to the report if they were to finish the text and suggested to the student to put it away somewhere and maybe they could listen to it someday.

The student teacher entered the classroom as the bell rang. Attendance check and distribution of test papers for approximately 20 students required six minutes. No comment was made about test results and no grades were announced. He then gave an extremely brief (less than five-minute) lecture on his topic. He appeared very uncertain of his notes and made no gesture to encourage the students to ask questions. Although this was to be an introduction to a unit, he made no such mention to the class and no follow-up was suggested.

The student teacher came late to his seventh grade social studies class. Before he called the class to order, he came back to chat with me. I asked him for the journal, but he replied that he had left it at home. After class, he told me that he had not started keeping the new cooperative section because neither he nor the supervising teacher knew how to use it. He was three days behind in his log. He seemed surprised that I should expect him to have lesson plans. After a prolonged session of taking roll, he opened his textbook and asked the class where they had been at the close of yesterday's lesson. The students argued about location, and the student resolved the matter by assigning a twenty-five minute study period during which the students were to read an assigned number of pages. After the study period, the student teacher checked upon their reading by improvising questions from the textbook.

Criterion III-Stimulating, Imaginative, Surgent vs. Dull, Routine, Unimaginative

Rating of 1-High

In a United States history class the student teacher had two boys who were especially gifted in drawing and not quite so gifted or motivated in history to prepare each day a comic-strip-type serial depicting the events which prompted the Civil War. The class contributed suggestions toward the end of the period.

The student teacher was observed in a speech class. As part of the introduction to a unit on story telling, the teacher told a story herself. The pupils were so interested in her enthusiastic presentation that they appeared to "hang" on her every word and gesture. At the conclusion they spontaneously broke into applause.

The student teacher in Latin was focusing on the development of vocabulary and the ability to converse in the language. He had a quick smile and good voice projection. He told a story in Latin and used overhead transparencies which he had prepared to illustrate key words. Whenever comprehension was obviously not achieved, he occasionally drew an illustration on the chalkboard or attempted to use humor to communicate. The students seemed to be enjoying this procedure and worked diligently at interpreting the story in

English. When he made a mistake in pronunciation, he smiled and said that teachers had a right to make mistakes, but students did not. All laughed.

Rating of 4-Midpoint

The industrial education student teacher spent a few minutes describing a part of a machine which would be disassembled in the lab. The description was presented in the form of a lecture with no use being made of the chalkboard or visual materials. He explained that the students would need to know this in order to work on their small engines. After the ten-minute presentation, he dismissed the students to the lab where they worked on their projects. He managed to give attention to all and kept the students busy on the project.

Student teacher in first year algebra opened the class asking if anyone wanted to see an example put on the board. He then worked several examples, putting them on the board in deliberate fashion. His presentation showed no appreciable verve, although he did smile and joke with the students occasionally.

In a bookkeeping class the student teacher consistently used examples which related to farming or rural type businesses which he was familiar with and was unaware of his failure to be stimulating until one of his "large city school" students asked him, "What is a combine?"

Rating of 7-Low

In a seventh grade English class a student remarked that he didn't see why they had to memorize the names and spelling of the Greek gods and goddesses. The student teacher replied that he never cared for this type of Literature either, but it was in the book and it was going to be on the test next Friday.

The student teacher in general business spent most of the time reading directly from the textbook. He read without expression and without any interruption to explain or to see if students had any response. The students sat quietly, but obviously were paying no attention. Some worked on other homework. Others looked through billfolds and other books. Many looked at the floor in obvious boredom.

The student teacher in American history class took a very long time to take roll visually, during which time the class became gradually quiet. As he was putting the attendance slip on the door, a messenger came with a note to be handed to a student. The student teacher accepted the note, mused over the name and only after several seconds did he call the name of the student. The student teacher waited silently until the student took the note, read it, went back to his seat for his books, and left the room. Eight minutes after the bell, he spoke for the first time to the class.

"Today we're starting a new unit on the westward movement. Does anyone have a question?" There was a long pause during which he gazed phlegmatically at the class and then down at his book. "Who was president at this time?" A student hurriedly leafed through his book and raised his hand. The student teacher sleepily looked at the hand a few seconds and called the student's name.

"Thomas Jefferson?"

"That's right, Thomas Jefferson. What were the dates of Jefferson's term in office?" There was another search through the pages by this same student while the student teacher patiently waited.

"1800 and 1804?"

"That's right, 1800 and 1804. But did he leave office in 1804?"

"No."

"That's right. He left office in 1808 after being elected in 1804."

The student teacher plodded through thirty minutes of this kind of discourse and then announced a study period. He sat at the desk during the study period. At the bell, he waited ten seconds before dismissing the class.

Criterion IV--Perceives Self as Competent vs. Perceives Self as Less than Adequate

Rating of 1--High

A student teacher in mathematics was assigned to teach a class in modern math for which the student teacher had little preparation. The student teacher informed the supervising teacher that after looking through the text he was sure he could do a good job with it and would present his lesson plans to the supervising teacher several days in advance so they could check them over carefully.

Student teacher was observed in a second-year French class. Before the class she told me that she was thoroughly enjoying student teaching and that she could hardly wait to begin each of her three daily classes. During the class, her pupils caught her in two errors. She, on both occasion, smiled, admitted her mistakes, thanked the student making the correction, and continued undisturbed with her presentation.

The student teacher in English III class was discussing a test with three or four students from the previous period when I entered the room. The students were complaining about the test, but the student teacher was smiling and unruffled. When he saw me, he instructed the students to see him at the sixth period and dismissed them. He shook my hand and asked where I would like to sit. He instructed a student to get a chair for me. As the student teacher began to speak to the class, the assistant principal came to the door and asked about a student. The student teacher in a clearly audible, informative voice said, "Oh, I gave him permission to do that. I didn't know that it was against the rules. The student was not at fault. I'm sorry. It won't happen again." He smiled at the assistant principal as he spoke. He looked directly at the students and waited for them to become quiet before he spoke. He interrupted students several times to insist that they speak more loudly. Several times he used "I think. . ." to summarize a point. After class, he explained that he had not written much in the journal because he did not have any problem.

Rating of 4--Midpoint

The student teacher in biology class taught a well organized, effective lesson. When I asked her how well she was progressing, she said, "All right, I guess." In the journal she expressed some anxiety about not being able to "reach" the students. She stated that she needed more experience and more background in subject matter. She said that she had had a difficult time in making out a test. In summary, she said, "Oh, I really think that I'm doing OK, but I sure have a lot to learn."

The student teacher appeared rather timid and had little to say. He listened in group conservations and asked few questions of anyone. In his class presentations, however, he appeared to know what he was talking about and projected his ideas in a confident manner. He was able to work individually with students in the class (drafting class) without showing any real timidity. He did not solicit criticisms from his supervisor, although he carefully listened to whatever he had to suggest by way of improvement.

The student teacher possessed good personal musicianship in illustrating techniques with the viola and piano but lacked security in directing the group. Her voice was submissive in direction. In counting the time it was difficult for students to hear. Finally, she went to the piano to play the accompaniment. The students could hear her play. She was at ease at this instrument and her directing was more successful.

Rating of 7--Low

The student teacher expressed concern that he did not know his content in algebra. When I reminded him that his index was better than a "B", he said he did not achieve that in courses he would be teaching. He stated that he feared that he would make mistakes in class and would embarrass himself in the presence of the students. He wanted to delay the beginning of his teaching and indicated that he had no idea of how to plan. He said that he could not see how simple arithmetic could be taught to the group that he would have to teach.

The student teacher was invited by the Future Teachers organization of the host school to meet with them and to talk with them about her university and the program leading to a degree in Education. The student teacher replied that she was much too busy since having several papers to grade and, besides, she wasn't very good at those sorts of things but if someone would contact the university maybe they would send someone out to do it. "I know, when my college supervisor comes around I'll mention it to him."

Student teacher was observed in a tenth-grade English class. During the period the class was somewhat noisy in that distractive "talking" was taking place between class members in different sections of the classroom. The student teacher made no effort to quiet these pupils. During our conference following this class, the student teacher said, "This is my worst class. They give all the teachers in the school trouble. I don't know what I am going to do with them. If I thought that I would ever get another class like this one, I don't think that I would enter teaching as a career."

Criterion V--Has Mastery of Facts and Organizing Principles of the Field vs. Has Only Minimum Knowledge of Field

Rating of 1-High

The student teacher in chemistry introduced the reaction of acids and bases. He called the attention of the students to various reactions which involve this phenomenon: baking soda and vinegar, battery acid and soda, erosion of teeth, acid soils, etc. He informed the students about the application of this principle which different local industries found useful. He also pointed out how the principle would affect their work in the study of chemistry. He answered the questions of the students with great poise and self-confidence. In conference after the class, the student teacher showed that he had a clear knowledge of the aims of his lesson. He knew how the principle would occur in future chemistry courses and had introduced it so that future application would be easy.

Student teacher was observed in a beginning biology class. In response to a pupil who asked why a certain relationship was true, the student teacher said, "The answer to this question is beyond the scope of this course and the information needed to answer it is gained in a second-year college chemistry course."

He continued, "However, since this is a college preparatory class, I feel that you will be able to partially understand the chemistry of the relationship if I diagram it on the chalkboard." He then briefly proved, using chemical symbols, the questioned relationship.

The objective of an eighth-grade English class was to discuss similes. Miss Banta read Part I from Evangeline. Her reading was excellent.

She paused to ask, "How can we decide Benedict was an old man?"

Helen repeated, "'He was hearty as an oak that is covered with snowflakes.'"

"'White as snow was his hair,'" said Rupert.

Rachel added, "'His cheeks were brown as oak leaves.'"

"Who can write on the board the names of these comparisons? Ruth?"

Ruth was pleased as she wrote similes.

"Do you agree, class?" asked Miss Banta. "Why are they similes?"

Rating of 4-Midpoint

A discussion of climate in a geography class was presented by Miss Tuttle. This particular school offered experience rich in laboratory potential with much equipment available. In this situation she was ill at ease. Trouble brewed when she was absorbed with the experiments of one student while the others took advantage. In conference she explained, "I am spending so many hours outside class in just learning about all this equipment. It is wonderful and I want to teach it this way. But, oh, my, I have to learn so much more."

The student teacher in algebra I introduced the working of simultaneous equations by the substitution method. When a student asked about the relative advantages of the addition-subtraction method as compared with the substitution method, the student teacher showed that he had not considered the question before. After some hesitancy, he answered the student satisfactorily. After class in conference, he indicated that he simply followed the book on such matters and felt no special obligation to demonstrate applicability of the principles involved.

Student teacher was observed in a business arithmetic class where the problems were concerned with insurance rates and premiums. He was extremely adept at explaining the correct solutions to the assigned problems but was vague when asked to explain certain types of insurance such as coinsurance, home owners insurance, endowment insurance, and disability insurance.

Rating of 7-Low

Student teacher was observed in a tenth-grade health class. The topic under discussion was communicable diseases--specifically, influenza, and other similar virus diseases. The student teacher followed the textbook topic by topic in leading the discussion. Even though approximately twenty percent of the pupils in his class were absent with a new virus disease--the "Hong Kong Flu"--the student teacher gave no indication of his awareness that such a disease existed.

Student teacher in social studies (economics) class presented a film-strip. After some delay because of a mistake in the previous rewinding of the film-strip, he began to show the panels with a simple statement that this was a film-strip. It was a commercial film-strip designed to show the operation and products of an industry. As he flipped to each new panel, the student teacher looked at the panel and then announced what was obvious to the students, "This is a pineapple field." The student teacher read laboriously the printed material while the students looked at the same material and read to themselves. At the end of the showing, a student asked how much of the film-strip would be included on the next test. The student teacher was vague and uncertain on this point. The period ended with the students wrangling about being held responsible for knowing such information. After class, the student teacher said that the only reason for showing the film-strip was that it was available.

In a health class students were expected to recite with text books open. The student teacher then outlined on the chalkboard the symptoms of various disorders as they were given in the text and students copied this into a notebook which was later to be turned in for a grade.

Criterion VI--Communicates Well and Empathetically vs. Communicates Ineffectively and Perfunctorily

Rating of 1--High

The mathematics student teacher reviewed the test results with the geometry class. As students raised questions, she sought to illustrate as vividly as possible. She explained the concept of perpendicular lines by using the desk to represent the surface and a yardstick to illustrate the concept. She constantly asked questions to get students to understand the mathematical process involved in each problem. She displayed a great amount of patience with students who had difficulty comprehending and not once did she appear irritated at the students for asking questions. She spoke loudly and clearly and she used examples which should have been familiar to the students.

In a class of auto mechanics the student teacher was teaching how to measure the cubic inches of displacement of the pistons. He used a cut away section of a model of a motor in his presentation and concluded by discussing with the students the difference between automobile models of their family autos such as "Dodge 383," "Chevrolet 307," etc.

Student teacher was observed in a beginning algebra class. As part of this lesson the pupils were first introduced to the division of polynomials. The student teacher gave the rules for and demonstrated the method of working such problems using literal numbers. She then asked the class if there were questions. A few hands were raised. She said, "I can see by your hands and the looks on some of your faces that you do not completely understand this new process. It is difficult and most pupils do have trouble with these problems at first. Perhaps, the process will become more clear to you if I work an example for you using arithmetical numbers in place of literal numbers." Following her second example, all the pupils were able to successfully begin working this type of problem.

Rating of 4--Midpoint

The student teacher was attempting to cause the students to think about literature and communications media by discussing current movies. He asked them to nominate their favorites for best movie, best actor, and best actress. He encouraged them to respond and attempted to create an informal environment by sitting on the desk. He did not (or could not) help them detect obvious personal interests (no real hard analysis of The Graduate as best movie, for example) and the main concepts of communication quality. There was no attempt to describe or demonstrate concepts of quality in motion-picture production. In other words, he did not get them to think beyond the emotional level. Some of his mannerisms, such as sitting on the desk for the entire period, may have proved distracting.

Mr. Stout began the lesson by making a study assignment for the new chapter in bookkeeping. He rapidly placed an example on the board and instructed the class to do the next three problems at their desks. Some of the students were working--others were puzzled.

Helen raised her hand saying, "I don't know where you want us to start."

The student teacher presented in lecture form the operation of the principle of diminishing returns. The student teacher's attitude was formal, and the students were taking notes. After twenty minutes of presentation, the student teacher asked, "Now, does everyone understand?" There was uneasy hesitance among the students, and one girl raised her hand.

"Question, Susan?"

"Well, I still don't understand how something can be worth more one time than another. How do you know which one to sell at a high price and which one to sell low?"

"Let's go through this again." The student teacher then repeated much of his original lecture. At the end of the second run, he asked Susan if she understood any better. Susan looked around the class.

"Well, I guess so."

"Does anyone else have a question?" No one asked any questions.

Rating of 7--Low

The student teacher in an eighth-grade English class was handing back a grammar test. Before he had completed handing back the test, several hands were up obviously in protest to the marking. Several students voiced their dissent.

"Last week you said it was all right to put a comma there."

"I have the same as Ellen, but you marked mine wrong."

"It's not fair."

"I don't see how you can expect us to learn this stuff when you don't tell us the same way twice."

The student teacher shouted for the students to become quiet and to ask their questions one at a time. As he tried to explain, however, other students would burst out with further objections. Finally, in exasperation, he gave the students a study period and threatened them with another test.

A student teacher in English kept asking students for parts of speech of sentences on the chalkboard. It was obvious that several students were confused and did not understand. The student teacher consistently replied to each wrong answer with, "No, now you know better than that," or "No, I'm ashamed of you for not knowing that."

The students in this typewriting class had just returned from a pep session. Miss Simpson had stayed in the room grading yesterday's papers. The papers were returned with no comments. Miss Simpson started the work with, "Now git (get) ready for the tin (ten) minute writings. We'll begin with the sintence (sentence) on line 12."

The room was noisy with students adjusting the typewriters. Henry couldn't get the margins adjusted. The bell kept ringing as he moved the carriage back and forth.

Miss Simpson, above the noise, snapped the clock and said, "Start."

About half the class started!

Criterion VII--Classroom Discourse Characterized by Reasoning or Creative Thinking vs. Classroom Discourse Characterized by Simple Recall

Rating of 1--High

The student teacher began his English class on communication by citing two books on communication which were of significance: Understanding Media and Human Use of Human Beings. He then briefly reviewed the elements of communication and the methods by which humans communicate.

The student teacher introduced the day's lesson by indicating that communication had two purposes: (1) to make people feel as you do, and (2) to get information. He then indicated that there are several ways of getting people to feel as you do:

1. Argument--"When I say, 'I like Kennedy; I think he should be President,' I am saying I think you should think so too."
2. Command--"If I ask you to open the window because it is hot in here, I say I think it is hot in here and you should too."
3. Observation--"When I say, 'It is a nice day outside,' I am saying I think it is a nice day outside and you should too."

He continued by outlining steps to follow in getting people to see things "our way." After this, he indicated how information is acquired by inductive and deductive processes.

The student teacher was observed in a twelfth-grade government class. The unit topic being studied by the class was "Individual Rights of American Citizens" and the specific topic under discussion was "Free Speech." The student teacher used a commercial recording which presented a simulated problem in this area. The problem was concerned with the "rights" involved if two newspapermen, using evidence based upon wiretapping and other means, were to print a story alleging graft on the part of a city alderman. The problem was open-ended. He asked each pupil to decide whether the story should be printed. Many pupils responded to the problem and a lively discussion ensued between those in favor of and those opposed to the printing of the story. During the discussion the role of the student teacher became mainly that of a moderator. Occasionally, he would have to re-focus the discussion upon the problem, but he did not attempt to instill his ideas upon the class. Rather, he allowed each individual pupil to study the facts and arrive at his own answer.

The student teacher was using the inductive method to develop the principle that individual freedom can be enhanced through social planning. In his government class, he asked the students to list things which students could do in 1900. The students listed such things as hikes in the woods, hunting, fishing, absence of high school, etc. Students then listed activities possible to them now. Mentioned were automobile driving, college, electric lights, telephones, plumbing, etc. The disadvantages of high taxes, congested facilities, curfews, etc., were compared with the disadvantages of few social facilities. The student teacher faced the students with the proposition that freedom is best defined as the power to do the greatest number of things. This was contrasted with the definition of freedom as absence of social restraint. The student teacher did not insist that the students take either position, at the moment, but that they consider which position they would choose later.

Rating of 4--Midpoint

The use of records of folk songs and ballads from the Uprising of 1645 was found in an English literature class.

"Mercenary soldiers would have what kind of emotional reactions that would differ from soldiers in the regular army?" quizzed Miss Thornberry. Then, "What kinds of feelings and emotions would you have if we were in a revolution today?"

Immediately, the responses grew more irresponsible with each student trying to outdo the last.The student teacher lost control of the class in the laughing and noise of the students as they heckled each other.

Student teacher: "What did the convict say that might frighten Pip?"

Student: "That he had a young man there who would eat out his heart and liver if he did not do what he said." (Class laughter)

Student teacher: "I know this sounds funny to you, but it might be very frightening to a small child. What do we have now which we sometimes tell children to frighten them?"

Student: "That the boogey man will get them."

Student teacher: "Right. And what did the convict demand of Pip?"

Student: "Some food and a file."

The class continued in this manner for about thirty minutes with the student teacher generally asking specific-recall questions, but occasionally asking a question which called for reasoning or expression of opinion.

The student teacher in Latin II class took most of the hour to hear the exercises which students had worked out previously. Following this, the students translated exercises from Latin to English and vice versa. In the translation exercises, the student teacher asked for alternative solutions and also posed the question, "If it had been this, what would it have meant?"

Rating of 7--Low

The student teacher in home economics, family living class, began the period with a buzz session which dealt with the problems of in-laws. After the buzz session, the class continued to give opinions and prejudices about in-laws. The student added some experiences of her own to illustrate that mothers-in-law try to dominate the family of their children. The student teacher then read a number of derogatory jokes and poems about mothers-in-law. After the period, she explained that her only purpose was to "talk over" the problems of in-laws.

Student teacher was observed in a United States history class. The Revolutionary War was the topic being studied. All questions asked by the student teacher during the period were similar to the following examples. "Who wrote 'Commonsense'?" "Who made the motion in the Second Continental Congress that the colonies declare their independence?" "Who made a famous midnight ride?" "What was the date of the Townsend Act?" "What was another name for the Tories?" "What river did Washington cross?"

A tenth-grade English grammar class was discussing participles. The teacher passed out duplicated practice sentences. Betty asked, "May we make up our own sentences?" The teacher commented, "No, you can't do it that way--Just fill in the blanks." Sentences irrelevant to the students' experiences were filled in with participles.

Criterion VIII--Directs Attention to Process of Thinking vs. Fails to Call Attention to Process of Thinking

Rating of 1--High

The student teacher asked the student to define a testimonial. After this was done, the following dialogue ensued.

Student teacher: "Can we rely on testimonials? I saw a commercial the other day featuring Andy Granatelli saying it was as difficult for him to get into a suit as it is to get into a racing car. So he found that he could rely on _______suits. Is that a reliable testimony?"

Student: "Don't they use famous people for commercials all the time?"

Student Teacher: "Yes, but is it reliable? Who is on the AAMCO Transmission commercial?"

Student: "Zsa Zsa Gabor."

Student teacher: "Is she an authority on transmissions?"

Student: "Not likely."

Student teacher: "Why is she probably not an authority? Why would she give a testimonial?"

Student: "She doesn't know anything about transmissions. They asked her to do it because she is famous."

Student teacher: "What kinds of testimonials can be trusted?"

The dialogue continued in this manner for some time.

In a world history class, the student teacher is searching for defensible generalizations to conclude a unit on Charlemagne. His questions were, "What can we say about Charlemagne's ascension to power?" The students noted such things as consolidating individual tribes into larger political units, preventing subdivision, strong sense of purpose. The student teacher then asked if these observations could apply to other rulers in such a way that a general principle could be formulated. The student teacher asked several times, "How valid would it be for us to state this as a principle?" At one point, he asked, "Can we say that it is a good policy to fight wars on several fronts simultaneously?" He ended the period with the statement that they would spend the next day trying to extract principles. His assignment was that they should consider the consequences of a domain based upon a strong personality which had been ruthlessly aggressive in pushing its objectives.

Student teacher was observed in a United States history class and the Reconstruction Era was being discussed. One part of the class period was devoted to a pupil debate--Resolved: That Congress had no right to impeach President Andrew Jackson. Following the debate, the student teacher summarized the arguments of the two sides, pointing out a faulty assumption on the part of one team and asking for evidence to support one statement made by the opposing team.

Rating of 4--Midpoint

A senior class in bookkeeping was involved in posting. The class was both presentation and laboratory. Each student was working on his own set. Miss Reynolds used the opaque projector for the ledger sheets. She asked,

"How do you get this closing entry?"

"To close it, what do we do?"

"Where do you want to begin this procedure?"

"Think about it for a moment."

"Do you see the difference?"

"Do you see what she did? Can you post it that way?"

Then as the teacher worked with individuals there was constant attention to the proper steps for the right answer. To one student she said pleasantly, "That was in the last chapter. Do you just learn the principles for one chapter then forget them? Use what you learned!" "Prove it!"

A class in economics was discussing supply and demand. A pupil commented that the prices always go up on new cars. The student teacher asked about prices on used cars. One boy said used convertibles cost more in the spring than in the fall. The class discussed why for a few minutes, then went back to the text. Nothing more was said about new cars.

"Did you outline your steps of procedure before you began?"

"I think I understand what it is you are having a problem with, but can you see that it will not affect the solution?"

Rating of 7--Low

Student teacher was observed in a United States history class. During a discussion of the consolidation of capital and labor the student teacher asked the following questions: "Does collective bargaining help the laboring men and not hurt the capitalists?" "Should our government have a right to compel capitalists and laborers to settle their difficulties by arbitration?" "How do strikes and boycotts affect the general public?" "Are trusts necessarily bad just because they are big?" "How may the laboring men win a strike?" After each question was asked, it was responded to orally by a class member. None of the pupil answers were commented on by the student teacher except by some expression of acknowledgement such as "O.K." or "All right." He would then proceed to another question. Students did not have any clear idea of what capitalists or laborers were. No terms were illustrated.

The student teacher in senior English class was leading a discussion of Chekhov's writings when a student burst out that he did not like his writing. "How do people decide who is great?" he asked. "Why is everyone so sure that Shakespeare is good?" The student teacher missed the chance to examine how opinions of classics are formed. He simply advised the student that since the textbook treated the works as important, the student should share the opinion. Throughout the remainder of the lesson, the student teacher accepted the textbook as the authority.

The following discussion occurred at the beginning of a government class:

Student teacher: "How many heard about the city council's passing a $500 raise for police and firemen? They sent it to the mayor and he vetoed it. Now they are trying to compromise. Last night they passed a resolution calling for a $200 raise. How many of you think our city policemen and firemen need a raise?" (Show of hands)
"How many opposed?" (Show of hands)

Student: "I don't think they need no raise. All they do is set around in restaurants and drink coffee all the time."

Second student: "I don't think so. One night I was stopped by one and he told me I was speedin'. I was only doin' 30. He had another friend in the car and they was hittin' the bottle, so I don't think they need a raise."

Third student: "Firemen need a raise before the police do."

Student teacher: "How many of you would be in favor of a $500 raise?" (Show of hands)
"$200" (Show of hands)
"We will see what happens."

The student teacher ignored cases of loose reasoning and resorted to a show of hands to determine a course of action.

APPENDIX D

SSTPP ASSESSMENTS																				TCAP DATA						
I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	CUM. GPA	EST. GPA	EST. X	EST. Y	EST. Z	C. C. T. T.	SAT VERB.	SAT MATH	MN	ML	P	R	D	LT	TP
1.000	.5143	.6177	.4356	.4615	.6106	.1632	.1516	.6486	.7774	.3398	.1750	-.0613	-.1166	.0768	.1747	.2551	-.0385	-.0362	.0221	-.0024	.1023	.0601	-.0584	.0358	-.1841	-.0834
564	1.000	.6162	.5378	.6335	.6041	.2150	.1987	.7322	.6910	.3336	.7015	-.1702	-.1595	.4148	.2681	.3208	-.0791	.0105	.0120	.1006	.0872	-.0011	.0053	.0249	-.2240	-.1041
564	564	1.000	.6461	.5720	.6959	.2368	.1707	.7400	.8813	.3664	.1979	-.1313	-.1267	.3002	.2305	.2937	-.0515	-.0372	-.0133	.0721	.1254	-.0217	-.0060	.0508	-.2410	-.1187
564	564	564	1.000	.5923	.5402	.1584	.1010	.6278	.7023	.1476	.5164	-.1631	-.0392	.4025	.2844	.3677	-.0513	.0739	.1148	.0752	.0468	-.0223	.0061	.0672	-.1860	-.0316
564	564	564	564	1.000	.6528	.2352	.2356	.7396	.6254	.4228	.7830	-.2562	-.1457	.4001	.2998	.3334	-.0146	.0140	.0350	.0587	.0650	.0262	-.0303	.0385	-.1696	-.0524
564	564	564	564	564	1.000	.2438	.2373	.7671	.8269	.4009	.3544	-.2467	-.2138	.3117	.3008	.3038	-.0553	-.0592	-.0380	.0324	.0794	.0681	.0086	-.0096	-.2219	-.1466
564	564	564	564	564	564	1.000	.6450	.2877	.2984	.9151	.2331	-.0026	-.1058	.0302	.0350	.1185	-.0346	.0268	-.0103	.0197	.0791	.0255	.0407	-.1000	-.1136	-.0559
564	564	564	564	564	564	564	1.000	.2522	.2428	.8903	.1788	-.0061	-.1056	.0817	.0433	.0536	.0211	.0019	-.0305	-.0076	.0266	.1494	-.1004	-.0423	-.0223	-.0385
564	564	564	564	564	564	564	564	1.000	.8180	.4852	.5894	-.2211	-.1851	.3716	.3704	.4414	-.0198	-.0770	-.0863	.0675	.1340	-.0022	.0135	-.0041	-.2461	-.1328
384	384	384	384	384	384	384	384	384	1.000	.1477	.3489	.1115	.1309	.3533	.3196	.3969	.0660	.0712	.0462	-.0717	-.1066	-.0408	-.0077	.0006	.2085	.0976
384	384	384	384	384	384	384	384	384	385	1.000	.1401	-.1235	-.2200	-.3860	.0794	.1606	-.0791	-.0237	-.1204	-.0258	.1598	.0848	.0187	.0103	-.3387	-.2722
384	384	384	384	384	384	384	384	384	385	385	1.000	-.1958	-.0976	.4848	.3121	.3709	-.0025	.0167	.0455	.0351	.0828	-.0606	-.0031	.0287	-.1009	.0010
539	539	539	539	539	539	539	539	539	365	365	365	1.000	.5935	-.5738	-.3670	-.1865	.3480	.3504	.2991	-.0830	.0236	-.0382	-.0602	.0262	.1220	-.0611
309	309	309	309	309	309	309	309	309	266	266	266	295	1.000	—	-.2694	-.1335	.3453	.5366	.5996	-.0228	.0041	-.0636	-.0570	.0240	.1394	.1286
52	52	52	52	52	52	52	52	52	35	35	35	51	0	1.000	.3766	.3321	—	-.1441	.0100	.2698	.1941	-.2874	.0001	-.0114	.0067	.1191
431	431	431	431	431	431	431	431	431	290	290	290	416	235	52	1.000	.3440	-.0649	-.1688	-.0660	.0073	-.0270	.0325	.0969	-.0566	-.0758	-.0468
534	534	534	534	534	534	534	534	534	367	367	367	510	292	52	426	1.000	-.1072	-.0437	-.0074	-.0217	.0033	[illegible]	.0384	-.0266	-.1024	-.0098
377	377	377	377	377	377	377	377	377	266	266	266	359	264	0	273	358	1.000	.1758	.1887	-.0935	-.0596	.088[illegible]	-.0057	-.0425	.0704	.0895
337	337	337	337	337	337	337	337	337	237	237	237	328	303	35	263	320	242	1.000	.6873	-.1704	.0956	.0399	-.2742	.3228	-.0161	.0734
337	337	337	337	337	337	337	337	337	237	237	237	328	303	35	263	320	242	337	1.000	-.1081	.0298	.0623	-.1861	.2216	.0267	.0473
445	445	445	445	445	445	445	445	445	316	316	316	426	253	36	333	423	316	267	267	1.000	-.1539	-.1144	.0146	-.0767	-.0574	-.1121
445	445	445	445	445	445	445	445	445	316	316	316	426	253	36	333	423	316	267	267	445	1.000	-.2911	-.3789	-.1868	-.2461	-.0951
445	445	445	445	445	445	445	445	445	316	316	316	426	253	36	333	423	316	267	267	445	445	1.000	-.3572	-.1585	.0571	-.0501
445	445	445	445	445	445	445	445	445	316	316	316	426	253	36	333	423	316	267	267	445	445	445	1.000	-.2977	-.2146	-.0612
445	445	445	445	445	445	445	445	445	316	316	316	426	253	36	333	423	316	267	267	445	445	445	445	1.000	-.0929	-.0893
445	445	445	445	445	445	445	445	445	316	316	316	426	253	36	333	423	316	267	267	445	445	445	445	445	1.000	.1500
445	445	445	445	445	445	445	445	445	316	316	316	426	253	36	333	423	316	267	267	445	445	445	445	445	445	1.000
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
426	426	426	426	426	426	426	426	426	299	299	299	407	299	0	308	401	499			424	424	424	424	424	424	424
424	424	424	424	424	424	424	424	424	298	298	298															
424	424	424	424	424	424	424	424	424	298	298	298															

MASTER MATRIX OF CORRELATION COEFFICIENTS

16 P.F. DATA

A	B	C	E	F	G	H	I	L	M	N	O	Q1	Q2	Q3	Q4	ANX.	EXTR.	MULT. R	F SCORE		
-.0230	-.0028	-.0970	-.0127	-.0390	-.0921	-.0894	-.0166	.0420	-.0097	.0656	.0051	-.0168	.0271	-.1114	.0573	.0993	-.0680	.2194	1.13	I	
.0323	-.0199	-.1564	-.0138	.0000	-.1617	-.0820	.0491	.0776	.0500	-.0351	.1140	-.0624	-.0159	-.1141	.1376	.1735	-.0342	.2254	1.20	II	
.0425	.0071	-.1608	-.1705	-.0733	-.0824	-.2368	-.0187	.0300	-.0481	.0110	.0568	-.0491	.0260	-.0715	.0784	.1439	-.1829	.3143	2.46	III	
.0301	.0067	-.1444	-.1677	-.1007	-.1588	-.2113	.1049	.0050	.0160	-.0017	.1541	-.0412	.0412	-.0834	.0685	.1663	-.1900	.3315	2.78	IV	SSTPP
.0418	-.0264	-.1345	-.0642	-.0412	-.1603	-.1251	.0059	.0497	.0165	.0396	.1081	-.0796	.0366	-.0971	.1166	.1515	-.0922	.2612	1.64	V	
-.0276	-.0462	-.1661	-.1731	-.0691	-.1097	-.1920	-.0240	.0441	.0187	.0331	.0954	-.0396	.0390	-.1122	.0981	.1649	-.1769	.2983	2.20	VI	
.1306	.0233	.0027	.0325	.1055	-.0535	.0782	-.0273	.0587	-.0426	.0305	.0168	-.0801	-.0835	-.0064	.0382	.0175	.1117	.2178	1.12	VII	
.0648	.0655	.0107	.0405	.1377	-.1371	.0828	-.0579	.0662	-.0522	.0904	.0163	-.0221	-.0332	-.0612	.0621	.0278	.1107	.2860	2.00	VIII	
-.0463	-.0402	-.1450	-.0858	-.0464	-.1322	-.1501	.0382	.0459	.0126	.0118	.1119	-.0743	.0222	-.1192	.1359	.1700	-.1082	.2620	1.66	IX	
.0139	.0229	.1366	.1420	.0898	.0886	.2279	.0010	-.0587	.0067	.0128	-.0469	.0567	-.0817	.1462	-.0973	-.1659	.1948	.2953	1.48	X	
.1724	.0239	-.0898	-.0557	.0514	-.0997	.0452	-.0535	.0117	.0698	.0972	.0146	-.0627	-.1086	-.0064	.0494	.0344	.0611	.3153	1.71	XI	
.0581	-.0337	-.0458	.0310	.0389	-.1279	.0242	.0806	.0502	.0979	-.0244	.0794	-.0943	-.0033	-.1255	.0766	.0794	.0402	.2401	0.94	XII	
-.0481	.2854	.0615	.1171	-.0145	.0482	-.0326	.0610	-.0161	.1121	.0267	-.0637	.0491	.1741	.0423	-.0146	-.0336	-.0174			CUM. GPA	
-.0417	.2092	.1138	.1680	-.0485	-.0726	-.0112	.0426	.0388	.1210	.0381	-.0234	.1379	.1137	-.0750	-.0897	-.0344	.0082			EST. GPA	
—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—			EST. X	
.0344	-.0800	-.0952	-.1288	-.0807	-.0544	-.1114	.0043	-.0372	-.0186	.0032	.0207	-.0542	-.0793	-.0816	.0801	.0896	.0993			EST. Y	
.0646	-.0471	-.0952	-.0257	.0064	-.1215	-.0619	.0387	-.0339	-.0240	-.0601	.0300	-.0264	-.0614	-.0283	.0569	.0635	-.0185			EST. Z	
-.0199	.2447	.0303	.0418	-.0863	-.0752	-.1290	.0971	-.0403	.0971	.0563	-.0444	.0782	.1556	-.0432	-.0568	-.0193	-.1041			CORNELL TEST	
																				SAT VERBAL	
																				SAT MATH	
-.0665	.0429	-.0504	-.0305	-.0331	-.0175	-.0438	-.0163	.0753	-.0699	-.1233	.0406	-.0421	-.0082	-.0061	.0151	.0496	-.0387			MN	
.0648	.0783	.0649	-.0018	.1179	.0796	.0772	-.0054	-.0679	-.0293	.0466	-.0768	.0056	-.0097	.0562	-.0477	-.0877	.0916			ML	TCAP
-.0814	-.0206	-.0558	-.0036	-.0376	-.0714	-.0148	.0184	-.0719	-.0109	.0469	-.0573	.0027	.0458	.0033	-.0159	-.0204	-.0493			P	
.0642	-.0463	-.0081	.0089	.0648	.0007	-.0077	.0347	.0645	-.0055	-.1675	.0608	-.0869	-.1013	.0127	.0764	.0622	.0509			R	
.0230	-.0071	-.0608	-.0475	-.0463	-.0124	-.0267	.0223	-.0082	-.0022	.0348	.0238	.0021	.1029	.0659	-.0645	-.0132	-.0546			D	
.0066	-.0020	.0554	.0034	-.0309	.1064	.0279	.0622	-.0458	-.0510	.0800	-.0681	.0137	-.0080	.0260	-.0797	-.0896	.0009			LT	
.0119	-.0387	.0684	.0462	-.0386	-.0059	-.0104	.0523	-.0431	-.0151	-.0798	-.0276	-.0952	.0761	-.0008	-.0461	-.0396	-.0165			TP	
1.000	.0124	.0506	.0804	.3330	.0446	.3152	.1655	-.0887	-.0094	.0276	.0237	-.0143	-.3117	.0199	-.0818	-.1145	.4634			A	
617	1.000	.0027	.0928	-.0234	-.0328	.0203	.0539	.0018	.0693	.0760	-.0962	.0888	.0559	-.0131	-.0555	-.0636	.0279			B	
617	617	1.000	.1011	.1588	.3195	.3118	-.1306	-.3461	-.1555	.2188	-.4499	.0566	-.0824	.2757	-.5377	-.6960	.2408			C	
617	617	617	1.000	.3806	-.1427	.3989	-.0199	.1428	.0694	-.0486	.0018	.1175	-.0896	-.2172	.1220	.0191	.6365			E	
617	617	617	617	1.000	-.0346	.4787	-.0952	.0880	-.1678	.0242	-.0136	.0291	-.4080	-.1088	.0243	-.0776	.8031			F	16 P.F.
617	617	617	617	617	1.000	.1235	-.1496	-.2710	.1894	.1869	-.2102	-.0964	-.1610	.3606	-.3021	-.3816	.0264			G	
617	617	617	617	617	617	1.000	-.0287	-.1437	-.0458	.0677	-.1585	.0611	-.3146	.1111	-.2265	-.4247	.8183			H	
617	617	617	617	617	617	617	1.000	.0150	.2254	-.0413	.0784	.0049	.1603	-.0271	.1033	.1090	-.0550			I	
617	617	617	617	617	617	617	617	1.000	.1698	-.1538	.3901	-.0217	.0211	-.3527	.4522	.6359	-.0069			L	
617	617	617	617	617	617	617	617	617	1.000	-.0450	.1361	.1082	.2465	-.0750	.1737	.1825	-.1089			M	
617	617	617	617	617	617	617	617	617	617	1.000	-.2327	.0173	-.0483	.2282	-.2266	-.2833	.0378			N	
617	617	617	617	617	617	617	617	617	617	617	1.000	-.0909	.0122	-.2993	.5286	.7463	-.0732			O	
617	617	617	617	617	617	617	617	617	617	617	617	1.000	.0962	-.0118	-.0499	-.0658	.0541			Q1	
617	617	617	617	617	617	617	617	617	617	617	617	617	1.000	-.0487	.0520	.1110	-.5059			Q2	
617	617	617	617	617	617	617	617	617	617	617	617	617	617	1.0000	-.3715	-.5669	-.0399			Q3	
617	617	617	617	617	617	617	617	617	617	617	617	617	617	617	1.000	.8327	-.0683			Q4	
617	617	617	617	617	617	617	617	617	617	617	617	617	617	617	617	1.000	-.2415			ANX.	
617	617	617	617	617	617	617	617	617	617	617	617	617	617	617	617	617	1.000			EXTR.	
																				MULT. R	
																				F-SCORE	

MASTER MATRIX OF CORRELATION COEFFICIENTS